

PCTWORLD INTELLECTUAL PROPERTY ORGANIZATION
International BureauDoc. Ref. **FP56 (1 of 7)**
Appl. No. 10/553,685

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|---|--------------------------|------------|---|----|------------|-------------------------|----|------------|-------------------------|----|------------|--------------------------|----|----|------------------|----------|--------------------------|----|------------------|----------|-------------------------|----|------------------|----------|-------------------------|----|------------------|----------|--------------------------|---|--|
| (51) International Patent Classification ⁷ : C12N 15/12, C07K 14/47, 16/18, G01N 33/566, C12Q 1/68, C12N 15/11, 15/62, A01K 67/027, A61K 38/00 | | A2 | (11) International Publication Number: WO 00/58473 (43) International Publication Date: 5 October 2000 (05.10.00) | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| (21) International Application Number: PCT/US00/08621 (22) International Filing Date: 31 March 2000 (31.03.00) (30) Priority Data: <table><tr><td>60/127,607</td><td>31 March 1999 (31.03.99)</td><td>US</td></tr><tr><td>60/127,636</td><td>2 April 1999 (02.04.99)</td><td>US</td></tr><tr><td>60/127,728</td><td>5 April 1999 (05.04.99)</td><td>US</td></tr><tr><td>09/540,763</td><td>30 March 2000 (30.03.00)</td><td>US</td></tr></table> (63) Related by Continuation (CON) or Continuation-in-Part (CIP) to Earlier Applications <table><tr><td>US</td><td>60/127,607 (CIP)</td></tr><tr><td>Filed on</td><td>31 March 1999 (31.03.99)</td></tr><tr><td>US</td><td>60/127,636 (CIP)</td></tr><tr><td>Filed on</td><td>2 April 1999 (02.04.99)</td></tr><tr><td>US</td><td>60/127,728 (CIP)</td></tr><tr><td>Filed on</td><td>5 April 1999 (05.04.99)</td></tr><tr><td>US</td><td>09/540,763 (CIP)</td></tr><tr><td>Filed on</td><td>30 March 2000 (30.03.00)</td></tr></table> (71) Applicant (for all designated States except US): CURAGEN CORPORATION [US/US]; 555 Long Wharf Drive, 11th Floor, New Haven, CT 06511 (US). | | 60/127,607 | 31 March 1999 (31.03.99) | US | 60/127,636 | 2 April 1999 (02.04.99) | US | 60/127,728 | 5 April 1999 (05.04.99) | US | 09/540,763 | 30 March 2000 (30.03.00) | US | US | 60/127,607 (CIP) | Filed on | 31 March 1999 (31.03.99) | US | 60/127,636 (CIP) | Filed on | 2 April 1999 (02.04.99) | US | 60/127,728 (CIP) | Filed on | 5 April 1999 (05.04.99) | US | 09/540,763 (CIP) | Filed on | 30 March 2000 (30.03.00) | (72) Inventors; and (75) Inventors/Applicants (for US only): SHIMKETS, Richard, A. [US/US]; 191 Leete Street, West Haven, CT 06516 (US). LEACH, Martin [GB/US]; 884 School Street, Webster, MA 01570 (US). (74) Agent: ELRIFI, Ivor, R.; Mintz, Levin, Cohn, Ferris, Glovsky and Popeo, P.C., One Financial Center, Boston, MA 02111 (US). (81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG). Published <i>Without international search report and to be republished upon receipt of that report.</i> | |
| 60/127,607 | 31 March 1999 (31.03.99) | US | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 60/127,636 | 2 April 1999 (02.04.99) | US | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 60/127,728 | 5 April 1999 (05.04.99) | US | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 09/540,763 | 30 March 2000 (30.03.00) | US | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| US | 60/127,607 (CIP) | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Filed on | 31 March 1999 (31.03.99) | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| US | 60/127,636 (CIP) | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Filed on | 2 April 1999 (02.04.99) | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| US | 60/127,728 (CIP) | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Filed on | 5 April 1999 (05.04.99) | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| US | 09/540,763 (CIP) | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Filed on | 30 March 2000 (30.03.00) | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| (54) Title: NUCLEIC ACIDS INCLUDING OPEN READING FRAMES ENCODING POLYPEPTIDES; "ORFX" | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| (57) Abstract <p>The present invention provides open reading frames ORFX, encoding isolated polypeptides, as well as polynucleotides encoding ORFX and antibodies that immunospecifically bind to ORFX or any derivative, variant, mutant, or fragment of the ORFX polypeptides, polynucleotides or antibodies. The invention additionally provides methods in which the ORFX polypeptide, polynucleotide and antibody are used in detection and treatment of a broad range of pathological states, as well as to other uses.</p> | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

| | | | | | | | |
|----|--------------------------|----|--|----|--|----|--------------------------|
| AL | Albania | ES | Spain | LS | Lesotho | SI | Slovenia |
| AM | Armenia | FI | Finland | LT | Lithuania | SK | Slovakia |
| AT | Austria | FR | France | LU | Luxembourg | SN | Senegal |
| AU | Australia | GA | Gabon | LV | Latvia | SZ | Swaziland |
| AZ | Azerbaijan | GB | United Kingdom | MC | Monaco | TD | Chad |
| BA | Bosnia and Herzegovina | GE | Georgia | MD | Republic of Moldova | TG | Togo |
| BB | Barbados | GH | Ghana | MG | Madagascar | TJ | Tajikistan |
| BE | Belgium | GN | Guinea | MK | The former Yugoslav Republic of Macedonia | TM | Turkmenistan |
| BF | Burkina Faso | GR | Greece | | | TR | Turkey |
| BG | Bulgaria | HU | Hungary | ML | Mali | TT | Trinidad and Tobago |
| BJ | Benin | IE | Ireland | MN | Mongolia | UA | Ukraine |
| BR | Brazil | IL | Israel | MR | Mauritania | UG | Uganda |
| BY | Belarus | IS | Iceland | MW | Malawi | US | United States of America |
| CA | Canada | IT | Italy | MX | Mexico | UZ | Uzbekistan |
| CF | Central African Republic | JP | Japan | NE | Niger | VN | Viet Nam |
| CG | Congo | KE | Kenya | NL | Netherlands | YU | Yugoslavia |
| CH | Switzerland | KG | Kyrgyzstan | NO | Norway | ZW | Zimbabwe |
| CI | Côte d'Ivoire | KP | Democratic People's Republic of Korea | NZ | New Zealand | | |
| CM | Cameroon | KR | Republic of Korea | PL | Poland | | |
| CN | China | KZ | Kazakhstan | PT | Portugal | | |
| CU | Cuba | LC | Saint Lucia | RO | Romania | | |
| CZ | Czech Republic | LI | Liechtenstein | RU | Russian Federation | | |
| DE | Germany | LK | Sri Lanka | SD | Sudan | | |
| DK | Denmark | LR | Liberia | SE | Sweden | | |
| EE | Estonia | | | SG | Singapore | | |

NOVEL POLYNUCLEOTIDES AND POLYPEPTIDES ENCODED THEREBY

5

BACKGROUND OF THE INVENTION

The invention relates generally to nucleic acids and polypeptides encoded thereby, and methods of using these nucleic acids and polypeptides.

10

SUMMARY OF THE INVENTION

The invention is based in part on the discovery of nucleic acids that include open reading frames encoding novel polypeptides, and on the polypeptides encoded thereby. The nucleic acids and polypeptides are collectively referred to herein as "ORFX".

Accordingly, in one aspect, the invention provides an isolated nucleic acid molecule (SEQ ID NO:2n-1, wherein n is an integer between 1-3161), that encodes novel polypeptide, or a fragment, homolog, analog or derivative thereof. The nucleic acid can include, *e.g.*, a nucleic acid sequence encoding a polypeptide at least 85% identical to a polypeptide comprising the amino acid sequences of SEQ ID NO:2n, wherein n is an integer between 1-3161. The nucleic acid can be, *e.g.*, a genomic DNA fragment, or a cDNA molecule.

Also included in the invention is a vector containing one or more of the nucleic acids described herein, and a cell containing the vectors or nucleic acids described herein.

The invention is also directed to host cells transformed with a recombinant expression vector comprising any of the nucleic acid molecules described above.

In another aspect, the invention includes a pharmaceutical composition that includes an ORFX nucleic acid and a pharmaceutically acceptable carrier or diluent.

In a further aspect, the invention includes a substantially purified ORF polypeptide, *e.g.*, any of the ORFX polypeptides encoded by an ORFX nucleic acid, and fragments, homologs, analogs, and derivatives thereof. The invention also includes a pharmaceutical composition that includes a ORFX polypeptide and a pharmaceutically acceptable carrier or diluent.

5 In a still a further aspect, the invention provides an antibody that binds specifically to an ORFX polypeptide. The antibody can be, *e.g.*, a monoclonal or polyclonal antibody, and fragments, homologs, analogs, and derivatives thereof. The invention also includes a pharmaceutical composition including ORFX antibody and a pharmaceutically acceptable carrier or diluent. The invention is also directed to isolated antibodies that bind to an epitope on a
10 polypeptide encoded by any of the nucleic acid molecules described above.

The invention also includes kits comprising any of the pharmaceutical compositions described above.

The invention further provides a method for producing an ORFX polypeptide by providing a cell containing a ORFX nucleic acid, *e.g.*, a vector that includes a ORFX nucleic
15 acid, and culturing the cell under conditions sufficient to express the ORFX polypeptide encoded by the nucleic acid. The expressed ORFX polypeptide is then recovered from the cell. Preferably, the cell produces little or no endogenous ORFX polypeptide. The cell can be, *e.g.*, a prokaryotic cell or eukaryotic cell.

The invention is also directed to methods of identifying an ORFX polypeptide or nucleic
20 acids in a sample by contacting the sample with a compound that specifically binds to the polypeptide or nucleic acid, and detecting complex formation, if present.

The invention further provides methods of identifying a compound that modulates the activity of a ORFX polypeptide by contacting ORFX polypeptide with a compound and determining whether the ORFX polypeptide activity is modified.

25 The invention is also directed to compounds that modulate ORFX polypeptide activity identified by contacting a ORFX polypeptide with the compound and determining whether the compound modifies activity of the ORFX polypeptide, binds to the ORFX polypeptide, or binds to a nucleic acid molecule encoding a ORFX polypeptide.

In a another aspect, the invention provides a method of determining the presence of or
30 predisposition of an ORFX-associated disorder in a subject. The method includes providing a sample from the subject and measuring the amount of ORFX polypeptide in the subject sample.

The amount of ORFX polypeptide in the subject sample is then compared to the amount of ORFX polypeptide in a control sample. An alteration in the amount of ORFX polypeptide in the subject protein sample relative to the amount of ORFX polypeptide in the control protein sample indicates the subject has a tissue proliferation-associated condition. A control sample is preferably taken from a matched individual, *i.e.*, an individual of similar age, sex, or other general condition but who is not suspected of having a tissue proliferation-associated condition. Alternatively, the control sample may be taken from the subject at a time when the subject is not suspected of having a tissue proliferation-associated disorder. In some embodiments, the ORFX is detected using a ORFX antibody.

In a further aspect, the invention provides a method of determining the presence of or predisposition of an ORFX-associated disorder in a subject. The method includes providing a nucleic acid sample, *e.g.*, RNA or DNA, or both, from the subject and measuring the amount of the ORFX nucleic acid in the subject nucleic acid sample. The amount of ORFX nucleic acid sample in the subject nucleic acid is then compared to the amount of an ORFX nucleic acid in a control sample. An alteration in the amount of ORFX nucleic acid in the sample relative to the amount of ORFX in the control sample indicates the subject has a tissue proliferation-associated disorder.

In a still further aspect, the invention provides method of treating or preventing or delaying a ORFX-associated disorder. The method includes administering to a subject in which such treatment or prevention or delay is desired a ORFX nucleic acid, a ORFX polypeptide, or an ORFX antibody in an amount sufficient to treat, prevent, or delay a tissue proliferation-associated disorder in the subject.

Unless otherwise defined, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. Although methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present invention, suitable methods and materials are described below. All publications, patent applications, patents, and other references mentioned herein are incorporated by reference in their entirety. In the case of conflict, the present specification, including definitions, will control. In addition, the materials, methods, and examples are illustrative only and not intended to be limiting.

Other features and advantages of the invention will be apparent from the following detailed description and claims.

DETAILED DESCRIPTION OF THE INVENTION

The invention provides novel polypeptides and nucleotides encoded thereby. The polynucleotides and their encoded polypeptides can be grouped according to the functions played by their gene products. Such functions include, structural proteins, proteins from which associated with metabolic pathways fatty acid metabolism, glycolysis, intermediary metabolism, calcium metabolism, proteases, and amino acid metabolism, etc.

Included in the invention are 3161 novel nucleic acid sequences and their encoded polypeptides. The sequences are collectively referred to as "ORFX nucleic acids" or ORFX polynucleotides" and the corresponding encoded polypeptide is referred to as a "ORFX polypeptide" or ORFX protein". For example, an ORFX nucleic acid according to the invention is a nucleic acid including an ORF1 nucleic acid, and an ORF polypeptide according to the invention is a polypeptide that includes the amino acid sequence of an ORF1 polypeptide. Unless indicated otherwise, "ORFX" is meant to refer to any of the ORF1-3161 sequences disclosed herein.

Table 1 provides a summary of the ORFX nucleic acids and their encoded polypeptides are summarized in Table 1. Nucleic acid sequences and polypeptide sequences for ORFX nucleic acids according to the invention is provided in the section of the specification entitled "Disclosed Sequences of ORFX Nucleic Acid and Polypeptide Sequences."

Column 1 of Table 1, entitled "ORF #", denotes an ORF number assigned to a nucleic acid containing an open reading frame according to the invention.

Column 2 of Table 1, entitled "Internal Identification number (Nucleic Acid Sequence Identification Number, Polypeptide Sequence Identification Number), provides an internal identification number for the indicated ORF, along with sequence identification numbers (SEQ ID NOs.) corresponding to the indicated ORF. In general, for an ORF_n according to the invention (wherein *n* is any integer from 1 to 3161), a nucleic acid corresponding to the ORF is SEQ ID NO:2*n*-1, and an amino acid sequence encoded by the ORF is SEQ ID NO:2*n*. For example, a nucleic acid sequence corresponding to an ORF1 nucleic acid is SEQ ID NO:1, and a polypeptide sequence corresponding to an ORF1 polypeptide is SEQ ID NO:2. Similarly, a

nucleic acid sequence corresponding to an ORF4 nucleic acid is SEQ ID NO:7, and a polypeptide sequence corresponding to an ORF4 polypeptide is SEQ ID NO:8; a nucleic acid sequence corresponding to an ORF198 nucleic acid sequence is SEQ ID NO:395, and a polypeptide sequence corresponding to an ORF198 polypeptide is SEQ ID NO:396. Nucleic acid sequences and polypeptide sequences for ORFX nucleic acids according to the invention are provided in the section of the specification entitled "Disclosed Sequences of ORFX Nucleic Acid and Polypeptide Sequences."

Column 2 of Table 1, entitled "Protein Similarity", lists previously described proteins that are related to polypeptides encoded by the ORFs. Genbank identifiers for the previously described proteins are provided. These can be retrieved from <http://www.ncbi.nlm.nih.gov/>.

To determine similarity to previously described proteins, polypeptides encoded by ORFX DNA sequences were tested using the Framesearch Algorithm against a nonredundant version of the GenPept Database from NCBI/Genbank. DNA sequences that had a score of '90' or above (Framesearch algorithm score, Edelman et. al. GCG Genetics) to a known protein were selected.

Open reading frames were extended beyond the region of the protein matched using standard DNA translation and codon tables. Novel proteins that lacked a protein match were translated against the standard genetic codons and proteins with an ORF at least 80 amino acids and containing a Methionine start are included in the Table.

Column 3 of Table 3, entitled "Protein Domains", lists previously described protein domains, designated by pfam entries, that are present in polypeptides encoded by the ORFs. Also included in column 3 are proteins in which these domains are present. The pfam entries can be retrieved from <http://pfam.wustl.edu/>. DNA sequences were translated in all six frames and tested using the Hmmer Algorithm against the Pfam Database (References to the algorithm and Pfam database can be found at <http://pfam.wustl.edu/>). Translated DNA sequences that matched a protein domain entry in the Pfam database AND had a score of '7.5' were selected.

Column 4 of Table 3, entitled "Protein Classification", lists the type of classification assigned for the protein, based on its homology. Examples of proteins in the classification include the following proteins:

Amylases

Amylase is responsible for endohydrolysis of 1,4-alpha-glucosidic linkages in oligosaccharides and polysaccharides. Variations in amylase gene may be indicative of delayed maturation and of various amylase producing neoplasms and carcinomas.

5 **Amyloid**

The serum amyloid A (SAA) proteins comprise a family of vertebrate proteins that associate predominantly with high density lipoproteins (HDL). The synthesis of certain members of the family is greatly increased in inflammation. Prolonged elevation of plasma SAA levels, as in chronic inflammation, 15 results in a pathological condition, called amyloidosis, which affects the liver, kidney and spleen and which is characterized by the highly insoluble accumulation of SAA in these tissues. Amyloid selectively inhibits insulin-stimulated glucose utilization and glycogen deposition in muscle, while not affecting adipocyte glucose metabolism. Deposition of fibrillar amyloid proteins intraneuronally, as neurofibrillary tangles, extracellularly, as plaques and in blood vessels, is characteristic of both Alzheimer's disease and aged Down's syndrome. 15 Amyloid deposition is also associated with type II diabetes mellitus.

Angiopoeitin

Members of the angiopoeitin/fibrinogen family have been shown to stimulate the generation of new blood vessels, inhibit the generation of new blood vessels, and perform several roles in blood clotting. This generation of new blood vessels, called angiogenesis, is also an 20 essential step in tumor growth in order for the tumor to get the blood supply it needs to expand. Variation in these genes may be predictive of any form of heart disease, numerous blood clotting disorders, stroke, hypertension and predisposition to tumor formation and metastasis. In particular, these variants may be predictive of the response to various antihypertensive drugs and chemotherapeutic and anti-tumor agents.

25 **Apoptosis-related proteins**

Active cell suicide (apoptosis) is induced by events such as growth factor withdrawal and toxins. It is controlled by regulators, which have either an inhibitory effect on programmed cell

death (anti-apoptotic) or block the protective effect of inhibitors (pro-apoptotic). Many viruses have found a way of countering defensive apoptosis by encoding their own anti-apoptosis genes preventing their target-cells from dying too soon. Variants of apoptosis related genes may be useful in formulation of anti-aging drugs.

5 **Cadherin, Cyclin, Polymerase, Oncogenes, Histones, Kinases**

Members of the cell division/cell cycle pathways such as cyclins, many transcription factors and kinases, DNA polymerases, histones, helicases and other oncogenes play a critical role in carcinogenesis where the uncontrolled proliferation of cells leads to tumor formation and eventually metastasis. Variation in these genes may be predictive of predisposition to any form
10 of cancer, from increased risk of tumor formation to increased rate of metastasis. In particular, these variants may be predictive of the response to various chemotherapeutic and anti-tumor agents.

Colony-stimulating factor-related proteins

Granulocyte/macrophage colony-stimulating factors are cytokines that act in
15 hematopoiesis by controlling the production, differentiation, and function of 2 related white cell populations of the blood, the granulocytes and the monocytes-macrophages.

Complement-related proteins

Complement proteins are immune associated cytotoxic agents, acting in a chain reaction to exterminate target cells to that were opsonized (primed) with antibodies, by forming a
20 membrane attack complex (MAC). The mechanism of killing is by opening pores in the target cell membrane. Variations in 20 complement genes or their inhibitors are associated with many autoimmune disorders. Modified serum levels of complement products cause edemas of various tissues, lupus (SLE), vasculitis, glomerulonephritis, renal failure, hemolytic anemia, thrombocytopenia, and arthritis. They interfere with mechanisms of ADCC (antibody dependent
25 cell cytotoxicity), severely impair immune competence and reduce phagocytic ability. Variants of complement genes may also be indicative of type I diabetes mellitus, meningitis neurological disorders such as nemaline myopathy, neonatal hypotonia, muscular disorders such as congenital myopathy and other diseases.

Cytochrome

The respiratory chain is a key biochemical pathway which is essential to all aerobic cells. There are five different cytochromes involved in the chain. These are heme bound proteins which serve as electron carriers. Modifications in these genes may be predictive of ataxia
5 areflexia, dementia and myopathic and neuropathic changes in muscles. Also, association with various types of solid tumors.

Kinesins

Kinesins are tubulin molecular motors that function to transport organelles within cells and to move chromosomes along microtubules during cell division. Modifications of these genes
10 may be indicative of neurological disorders such as Pick disease of the brain, tuberous sclerosis.

Cytokines, Interferon, Interleukin

Members of the cytokine families are known for their potent ability to stimulate cell growth and division even at low concentrations. Cytokines such as erythropoietin are cell-specific in their growth stimulation; erythropoietin is useful for the stimulation of the
15 proliferation of erythroblasts. Variants in cytokines may be predictive for a wide variety of diseases, including cancer predisposition.

G-protein coupled receptors

G-protein coupled receptors (also called R7G) are an extensive group of hormones, neurotransmitters, odorants and light receptors which transduce extracellular signals by
20 interaction with guanine nucleotide-binding (G) proteins. Alterations in genes coding for G-coupled proteins may be involved in and indicative of a vast number of physiological conditions. These include blood pressure regulation, renal dysfunctions, male infertility, dopamine associated cognitive, emotional, and endocrine functions, hypercalcemia, chondrodysplasia and osteoporosis, pseudohypoparathyroidism, growth retardation and
25 dwarfism.

Thioesterases

Eukaryotic thiol proteases are a family of proteolytic enzymes which contain an active site cysteine. Catalysis proceeds through a thioester intermediate and is facilitated by a nearby histidine side chain; an asparagine completes the essential catalytic triad. Variants of thioester associated genes may be predictive of neuronal disorders and mental illnesses such as Ceroid Lipoffiscinosis, Neuronal 1, Infantile, Santavuori disease and more.

The key to the molecule type is as follows:

| 10 | Abbrev: | Title: |
|----|-------------------|--|
| | amylase | amylase protein |
| | amylaseinhib | amylase inhibitor |
| | amyloid | amyloid protein |
| 15 | apoptosis | apoptosis associated protein |
| | apoptosisinhib | apoptosis inhibitors |
| | apoptosisrecep | apoptosis receptors |
| | ATPase_associated | ATPase associated protein |
| | biotindp | biotin dependent enzyme/protein |
| 20 | cadherin | cadherin protein |
| | calcium_channel | calcium channel protein |
| | carboxylase | carboxylase protein |
| | cathepsin | cathepsin/carboxypeptidases |
| | cathepsininhib | cathepsin/carboxypeptidase inhibitor |
| 25 | chloride_channel | chloride channel protein |
| | collagen | collagen |
| | complement | complement protein |
| | complementrecept | complement receptor protein |
| | complementinhib | complement inhibitor |
| 30 | csf | colony stimulating factor |
| | csfrecept | colony stimulating factor receptor |
| | cyclin | cyclin protein |
| | cyto450 | cytochrome p450 protein |
| | cytochrome | cytochrome related protein |
| 35 | deaminase | deaminase |
| | dehydrogenase | dehydrogenase |
| | desaturase | desaturase |
| | dna_ma_bind | DNA/RNA binding protein/factor |
| | dna_ma_inhib | DNA/RNA binding protein/factor inhibitor |
| 40 | dynein | dynein |

| | | |
|----|--------------------|-----------------------------------|
| | elastase | elastase |
| | elastaseinhib | elastase inhibitor |
| | eph | EPH family of tyrosine kinases |
| | esterase | esterase |
| 5 | esteraseinhib | esterase inhibitor |
| | fgf | fibroblast growth factor |
| | fgfreceptor | fibroblast growth factor receptor |
| | gaba | GABA receptor |
| | glucoamylase | glucoamylase |
| 10 | glucoronidase | glucoronidase |
| | glycoprotein | glycoprotein |
| | Guanylyl | guanylylate cyclase |
| | helicase | helicase |
| | histone | histone |
| 15 | HOM | homologous |
| | homeobox | homeobox protein |
| | hydrolase | hydrolase |
| | hydroxysteroid | hydroxysteroid associated protein |
| | hypoxanthine | hypoxanthine associated protein |
| 20 | immunoglob | immunoglobulin |
| | immunoglobrecept | immunoglobulin receptor |
| | interferon | interferon |
| | interleukin | interleukin |
| | interleukinrecept | interleukin receptor |
| 25 | isomerase | isomerase |
| | isomeraseinhibitor | isomerase inhibitor |
| | isomerasereceptor | isomerase receptor |
| | kinase | kinase |
| | kinaseinhibitor | kinase inhibitor |
| 30 | kinasereceptor | kinase receptor |
| | kinesin | kinesin |
| | laminin | laminin associated protein |
| | lipase | lipase |
| | metallothionein | metallothionein |
| 35 | MHC | major histocompatibility complex |
| | misc_channel | miscellaneous channel |
| | ngf | nerve growth factor |
| | nuci_recpt | nuclear receptor |
| | nuclease | nuclease |
| 40 | oncogene | oncogene associated protein |
| | oxidase | oxidase |
| | oxygenase | oxygenase |
| | peptidase | peptidase |
| | peroxidase | peroxidase |
| 45 | phosphatase | phosphatase |
| | phosphataseinhib | phosphatase inhibitor |

| | | |
|----|--|---|
| | phosphorylase PIR | phosphorylase PIR DATABASE (release 56, 29-OCT-1998) |
| 5 | polymerase potassium_channel prostaglandin protease proteaseinhib reductase | polymerase potassium channel protein prostaglandin protease protease inhibitor reductase |
| 10 | ribosomalprot RTR | ribosomal associated protein EMBLDATABASE translated entries not to be incorporated into SWISS-PROT (20-JUL-1998) |
| 15 | SIM SPTR | similar EMBL DATABASE translated entries to be incorporated into SWISS-PROT (20-JUL-1998) |
| 20 | struct sulfotransferase SWP | structural associated protein sulfotransferase SWISS-PROT DATABASE (release 18-OCT-1998) |
| 25 | SWPN synthase tgf tgfreceptor thioesterase thiolase tm7 | SWISS-PROT Update (release 11-NOV-98) synthase transforming growth factor transforming growth factor receptor thioesterase thiolase seven transmembrane domain G-protein coupled receptor |
| 30 | tnf traffic tnfreceptor TRN | necrosis factor receptor tumor necrosis factor tumor trafficking associated protein EMBL DATABASE translated entries update (20-JUL-1998) |
| 35 | transcriptfactor transferase transport tubulin ubiquitin | transcription factor transferase transport protein tubulin ubiquitin |
| 40 | unclassified water channel | Protein not categorized into one of the aforementioned protein families water channel protein |

Column 5 of Table 1, entitled, "Cells or Tissues in Which Gene is Expressed", denotes tissues, represented by five digit numbers, in which RNA homologous to the ORF nucleic acid sequences is present. Tissues or cells corresponding to the numbers are provided in Table 2.

ORFX nucleic acids, and their encoded polypeptides, according to the invention are useful in a variety of applications and contexts. For example, various ORFX nucleic acids and polypeptides according to the invention are useful, *inter alia*, as novel members of the protein families indicated in Table 1, and/or according to the presence of domains and sequence relatedness to previously described proteins as summarized in Table 1.

ORFX nucleic acids and polypeptides according to the invention can also be used to identify cell types listed in Table 1 for an indicated ORFX according to the invention. Additional utilities for ORFX nucleic acids and polypeptides according to the invention are disclosed herein.

ORFX Nucleic Acids

The novel nucleic acids of the invention include those that encode an ORFX or ORFX-like protein, or biologically active portions thereof. The nucleic acids include nucleic acids encoding polypeptides that include the amino acid sequence of one or more of SEQ ID NO:2*n*, wherein *n* = 1 to 3161. The encoded polypeptides can thus include, *e.g.*, the amino acid sequences of SEQ ID NO: 2, 4, 6, 8, 10, . . . , 6310, 6312, 6314, 6316, 6318, 6320, and/or 6322.

In some embodiments, a nucleic acid encoding a polypeptide having the amino acid sequence of one or more of SEQ ID NO:2*n* (wherein *n* = 1 to 3161) includes the nucleic acid sequence of any of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161), or a fragment thereof. Additionally, the invention includes mutant or variant nucleic acids of any of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161), or a fragment thereof, any of whose bases may be changed from the disclosed sequence while still encoding a protein that maintains its ORFX-like activities and physiological functions. The invention further includes the complement of the nucleic acid sequence of any of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161), including fragments, derivatives,

analog and homolog thereof. The invention additionally includes nucleic acids or nucleic acid fragments, or complements thereto, whose structures include chemical modifications.

Also included are nucleic acid fragments sufficient for use as hybridization probes to identify ORFX-encoding nucleic acids (*e.g.*, ORFX mRNA) and fragments for use as
5 polymerase chain reaction (PCR) primers for the amplification or mutation of ORFX nucleic acid molecules. As used herein, the term "nucleic acid molecule" is intended to include DNA molecules (*e.g.*, cDNA or genomic DNA), RNA molecules (*e.g.*, mRNA), analogs of the DNA or RNA generated using nucleotide analogs, and derivatives, fragments and homologs thereof. The nucleic acid molecule can be single-stranded or double-stranded, but preferably is
10 double-stranded DNA.

"Probes" refer to nucleic acid sequences of variable length, preferably between at least about 10 nucleotides (nt), 100 nt, or as many as about, *e.g.*, 6,000 nt, depending on use. Probes are used in the detection of identical, similar, or complementary nucleic acid sequences. Longer length probes are usually obtained from a natural or recombinant source, are highly specific and
15 much slower to hybridize than oligomers. Probes may be single- or double-stranded and designed to have specificity in PCR, membrane-based hybridization technologies, or ELISA-like technologies.

An "isolated" nucleic acid molecule is one that is separated from other nucleic acid molecules that are present in the natural source of the nucleic acid. Examples of isolated nucleic
20 acid molecules include, but are not limited to, recombinant DNA molecules contained in a vector, recombinant DNA molecules maintained in a heterologous host cell, partially or substantially purified nucleic acid molecules, and synthetic DNA or RNA molecules. Preferably, an "isolated" nucleic acid is free of sequences which naturally flank the nucleic acid (*i.e.*, sequences located at the 5' and 3' ends of the nucleic acid) in the genomic DNA of the organism
25 from which the nucleic acid is derived. For example, in various embodiments, the isolated ORFX nucleic acid molecule can contain less than about 50 kb, 25 kb, 5 kb, 4 kb, 3 kb, 2 kb, 1 kb, 0.5 kb or 0.1 kb of nucleotide sequences which naturally flank the nucleic acid molecule in genomic DNA of the cell from which the nucleic acid is derived. Moreover, an "isolated" nucleic acid molecule, such as a cDNA molecule, can be substantially free of other cellular
30 material or culture medium when produced by recombinant techniques, or of chemical precursors or other chemicals when chemically synthesized.

A nucleic acid molecule of the present invention, *e.g.*, a nucleic acid molecule having the nucleotide sequence of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161), or a complement of any of this nucleotide sequence, can be isolated using standard molecular biology techniques and the sequence information provided herein. Using all or a portion of the nucleic acid sequence of any of SEQ ID NO:2*n*-1 (wherein *n*=1 to 3161) as a hybridization probe, ORFX nucleic acid sequences can be isolated using standard hybridization and cloning techniques (*e.g.*, as described in Sambrook *et al.*, eds., MOLECULAR CLONING: A LABORATORY MANUAL 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989; and Ausubel, *et al.*, eds., CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, New York, NY, 1993.)

A nucleic acid of the invention can be amplified using cDNA, mRNA or alternatively, genomic DNA, as a template and appropriate oligonucleotide primers according to standard PCR amplification techniques. The nucleic acid so amplified can be cloned into an appropriate vector and characterized by DNA sequence analysis. Furthermore, oligonucleotides corresponding to ORFX nucleotide sequences can be prepared by standard synthetic techniques, *e.g.*, using an automated DNA synthesizer.

As used herein, the term "oligonucleotide" refers to a series of linked nucleotide residues, which oligonucleotide has a sufficient number of nucleotide bases to be used in a PCR reaction. A short oligonucleotide sequence may be based on, or designed from, a genomic or cDNA sequence and is used to amplify, confirm, or reveal the presence of an identical, similar or complementary DNA or RNA in a particular cell or tissue. Oligonucleotides comprise portions of a nucleic acid sequence having about 10 nt, 50 nt, or 100 nt in length, preferably about 15 nt to 30 nt in length. In one embodiment, an oligonucleotide comprising a nucleic acid molecule less than 100 nt in length would further comprise at least 6 contiguous nucleotides of any of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161), or a complement thereof. Oligonucleotides may be chemically synthesized and may be used as probes.

In another embodiment, an isolated nucleic acid molecule of the invention comprises a nucleic acid molecule that is a complement of the nucleotide sequence shown in any of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161). In another embodiment, an isolated nucleic acid molecule of the invention comprises a nucleic acid molecule that is a complement of the nucleotide sequence shown in any of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161), or a portion of this nucleotide sequence. A nucleic acid molecule that is complementary to the nucleotide sequence shown in

is one that is sufficiently complementary to the nucleotide sequence shown in of any of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161) that it can hydrogen bond with little or no mismatches to the nucleotide sequence shown in of any of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161), thereby forming a stable duplex.

- 5 As used herein, the term "complementary" refers to Watson-Crick or Hoogsteen base pairing between nucleotides units of a nucleic acid molecule, and the term "binding" means the physical or chemical interaction between two polypeptides or compounds or associated polypeptides or compounds or combinations thereof. Binding includes ionic, non-ionic, Von der Waals, hydrophobic interactions, etc. A physical interaction can be either direct or indirect.
- 10 Indirect interactions may be through or due to the effects of another polypeptide or compound. Direct binding refers to interactions that do not take place through, or due to, the effect of another polypeptide or compound, but instead are without other substantial chemical intermediates.

- Moreover, the nucleic acid molecule of the invention can comprise only a portion of the nucleic acid sequence of any of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161), *e.g.*, a fragment that can be used as a probe or primer, or a fragment encoding a biologically active portion of ORFX. Fragments provided herein are defined as sequences of at least 6 (contiguous) nucleic acids or at least 4 (contiguous) amino acids, a length sufficient to allow for specific hybridization in the case of nucleic acids or for specific recognition of an epitope in the case of amino acids, respectively,
- 20 and are at most some portion less than a full length sequence. Fragments may be derived from any contiguous portion of a nucleic acid or amino acid sequence of choice. Derivatives are nucleic acid sequences or amino acid sequences formed from the native compounds either directly or by modification or partial substitution. Analogs are nucleic acid sequences or amino acid sequences that have a structure similar to, but not identical to, the native compound but
- 25 differs from it in respect to certain components or side chains. Analogs may be synthetic or from a different evolutionary origin and may have a similar or opposite metabolic activity compared to wild type.

- Derivatives and analogs may be full length or other than full length, if the derivative or analog contains a modified nucleic acid or amino acid, as described below. Derivatives or
- 30 analogs of the nucleic acids or proteins of the invention include, but are not limited to, molecules comprising regions that are substantially homologous to the nucleic acids or proteins of the

invention, in various embodiments, by at least about 70%, 80%, 85%, 90%, 95%, 98%, or even 99% identity (with a preferred identity of 80-99%) over a nucleic acid or amino acid sequence of identical size or when compared to an aligned sequence in which the alignment is done by a computer homology program known in the art, or whose encoding nucleic acid is capable of hybridizing to the complement of a sequence encoding the aforementioned proteins under stringent, moderately stringent, or low stringent conditions. See *e.g.* Ausubel, *et al.*, CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, New York, NY, 1993, and below. An exemplary program is the Gap program (Wisconsin Sequence Analysis Package, Version 8 for UNIX, Genetics Computer Group, University Research Park, Madison, WI) using the default settings, which uses the algorithm of Smith and Waterman (Adv. Appl. Math., 1981, 2: 482-489, which is incorporated herein by reference in its entirety).

A "homologous nucleic acid sequence" or "homologous amino acid sequence," or variations thereof, refer to sequences characterized by a homology at the nucleotide level or amino acid level as discussed above. Homologous nucleotide sequences encode those sequences coding for isoforms of ORFX polypeptide. Isoforms can be expressed in different tissues of the same organism as a result of, for example, alternative splicing of RNA. Alternatively, isoforms can be encoded by different genes. In the present invention, homologous nucleotide sequences include nucleotide sequences encoding for a ORFX polypeptide of species other than humans, including, but not limited to, mammals, and thus can include, *e.g.*, mouse, rat, rabbit, dog, cat, cow, horse, and other organisms. Homologous nucleotide sequences also include, but are not limited to, naturally occurring allelic variations and mutations of the nucleotide sequences set forth herein. A homologous nucleotide sequence does not, however, include the nucleotide sequence encoding human ORFX protein. Homologous nucleic acid sequences include those nucleic acid sequences that encode conservative amino acid substitutions (see below) in any of SEQ ID NO:2*n* (wherein *n* = 1 to 3161) as well as a polypeptide having ORFX activity. Biological activities of the ORFX proteins are described below. A homologous amino acid sequence does not encode the amino acid sequence of a human ORFX polypeptide.

The nucleotide sequence determined from the cloning of the human ORFX gene allows for the generation of probes and primers designed for use in identifying the cell types disclosed and/or cloning ORFX homologues in other cell types, *e.g.*, from other tissues, as well as ORFX homologues from other mammals. The probe/primer typically comprises a substantially purified

oligonucleotide. The oligonucleotide typically comprises a region of nucleotide sequence that hybridizes under stringent conditions to at least about 12, 25, 50, 100, 150, 200, 250, 300, 350 or 400 or more consecutive sense strand nucleotide sequence of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161); or an anti-sense strand nucleotide sequence of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161); or of a naturally occurring mutant of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161).

Probes based on the human ORFX nucleotide sequence can be used to detect transcripts or genomic sequences encoding the same or homologous proteins. In various embodiments, the probe further comprises a label group attached thereto, *e.g.*, the label group can be a radioisotope, a fluorescent compound, an enzyme, or an enzyme co-factor. Such probes can be used as a part of a diagnostic test kit for identifying cells or tissue which misexpress a ORFX protein, such as by measuring a level of a ORFX-encoding nucleic acid in a sample of cells from a subject *e.g.*, detecting ORFX mRNA levels or determining whether a genomic ORFX gene has been mutated or deleted.

"A polypeptide having a biologically active portion of ORFX" refers to polypeptides exhibiting activity similar, but not necessarily identical to, an activity of a polypeptide of the present invention, including mature forms, as measured in a particular biological assay, with or without dose dependency. A nucleic acid fragment encoding a "biologically active portion of ORFX" can be prepared by isolating a portion of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161), that encodes a polypeptide having a ORFX biological activity (biological activities of the ORFX proteins are summarized in Table 1), expressing the encoded portion of ORFX protein (*e.g.*, by recombinant expression *in vitro*) and assessing the activity of the encoded portion of ORFX. For example, a nucleic acid fragment encoding a biologically active portion of ORFX can optionally include a domain as shown in Table 1, column 4.

ORFX variants

The invention further encompasses nucleic acid molecules that differ from the disclosed ORFX nucleotide sequences due to degeneracy of the genetic code. These nucleic acids thus encode the same ORFX protein as that encoded by the nucleotide sequence shown in SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161). In another embodiment, an isolated nucleic acid molecule of the invention has a nucleotide sequence encoding a protein having an amino acid sequence shown in any of SEQ ID NO:2*n* (wherein *n* = 1 to 3161).

In addition to the human ORFX nucleotide sequence shown in any of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161), it will be appreciated by those skilled in the art that DNA sequence polymorphisms that lead to changes in the amino acid sequences of ORFX may exist within a population (*e.g.*, the human population). Such genetic polymorphism in the ORFX gene may exist among individuals within a population due to natural allelic variation. As used herein, the terms "gene" and "recombinant gene" refer to nucleic acid molecules comprising an open reading frame encoding a ORFX protein, preferably a mammalian ORFX protein. Such natural allelic variations can typically result in 1-5% variance in the nucleotide sequence of the ORFX gene. Any and all such nucleotide variations and resulting amino acid polymorphisms in ORFX that are the result of natural allelic variation and that do not alter the functional activity of ORFX are intended to be within the scope of the invention.

Moreover, nucleic acid molecules encoding ORFX proteins from other species, and thus that have a nucleotide sequence that differs from the human sequence of any of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161), are intended to be within the scope of the invention. Nucleic acid molecules corresponding to natural allelic variants and homologues of the ORFX cDNAs of the invention can be isolated based on their homology to the human ORFX nucleic acids disclosed herein using the human cDNAs, or a portion thereof, as a hybridization probe according to standard hybridization techniques under stringent hybridization conditions.

In another embodiment, an isolated nucleic acid molecule of the invention is at least 6 nucleotides in length and hybridizes under stringent conditions to the nucleic acid molecule comprising the nucleotide sequence of any of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161). In another embodiment, the nucleic acid is at least 10, 25, 50, 100, 250, 500 or 750 nucleotides in length. In another embodiment, an isolated nucleic acid molecule of the invention hybridizes to the coding region. As used herein, the term "hybridizes under stringent conditions" is intended to describe conditions for hybridization and washing under which nucleotide sequences at least 60% homologous to each other typically remain hybridized to each other.

Homologs (*i.e.*, nucleic acids encoding ORFX proteins derived from species other than human) or other related sequences (*e.g.*, paralogs) can be obtained by low, moderate or high stringency hybridization with all or a portion of the particular human sequence as a probe using methods well known in the art for nucleic acid hybridization and cloning.

As used herein, the phrase "stringent hybridization conditions" refers to conditions under which a probe, primer or oligonucleotide will hybridize to its target sequence, but to no other sequences. Stringent conditions are sequence-dependent and will be different in different circumstances. Longer sequences hybridize specifically at higher temperatures than shorter sequences. Generally, stringent conditions are selected to be about 5°C lower than the thermal melting point (T_m) for the specific sequence at a defined ionic strength and pH. The T_m is the temperature (under defined ionic strength, pH and nucleic acid concentration) at which 50% of the probes complementary to the target sequence hybridize to the target sequence at equilibrium. Since the target sequences are generally present at excess, at T_m , 50% of the probes are occupied at equilibrium. Typically, stringent conditions will be those in which the salt concentration is less than about 1.0 M sodium ion, typically about 0.01 to 1.0 M sodium ion (or other salts) at pH 7.0 to 8.3 and the temperature is at least about 30°C for short probes, primers or oligonucleotides (*e.g.*, 10 nt to 50 nt) and at least about 60°C for longer probes, primers and oligonucleotides. Stringent conditions may also be achieved with the addition of destabilizing agents, such as formamide.

Stringent conditions are known to those skilled in the art and can be found in CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, N.Y. (1989), 6.3.1-6.3.6. Preferably, the conditions are such that sequences at least about 65%, 70%, 75%, 85%, 90%, 95%, 98%, or 99% homologous to each other typically remain hybridized to each other. A non-limiting example of stringent hybridization conditions is hybridization in a high salt buffer comprising 6X SSC, 50 mM Tris-HCl (pH 7.5), 1 mM EDTA, 0.02% PVP, 0.02% Ficoll, 0.02% BSA, and 500 mg/ml denatured salmon sperm DNA at 65°C. This hybridization is followed by one or more washes in 0.2X SSC, 0.01% BSA at 50°C. An isolated nucleic acid molecule of the invention that hybridizes under stringent conditions to the sequence of any of SEQ ID NO:2*n*-1 (wherein $n = 1$ to 3161) corresponds to a naturally occurring nucleic acid molecule. As used herein, a "naturally-occurring" nucleic acid molecule refers to an RNA or DNA molecule having a nucleotide sequence that occurs in nature (*e.g.*, encodes a natural protein).

In a second embodiment, a nucleic acid sequence that is hybridizable to the nucleic acid molecule comprising the nucleotide sequence of any of SEQ ID NO:2*n*-1 (wherein $n = 1$ to 3161), or fragments, analogs or derivatives thereof, under conditions of moderate stringency is provided. A non-limiting example of moderate stringency hybridization conditions are

hybridization in 6X SSC, 5X Denhardt's solution, 0.5% SDS and 100 mg/ml denatured salmon sperm DNA at 55°C, followed by one or more washes in 1X SSC, 0.1% SDS at 37°C. Other conditions of moderate stringency that may be used are well known in the art. See, e.g., Ausubel *et al.* (eds.), 1993, CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, NY, and Kriegler, 1990, GENE TRANSFER AND EXPRESSION, A LABORATORY MANUAL, Stockton Press, NY.

In a third embodiment, a nucleic acid that is hybridizable to the nucleic acid molecule comprising the nucleotide sequence of any of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161), or fragments, analogs or derivatives thereof, under conditions of low stringency, is provided. A non-limiting example of low stringency hybridization conditions are hybridization in 35% formamide, 5X SSC, 50 mM Tris-HCl (pH 7.5), 5 mM EDTA, 0.02% PVP, 0.02% Ficoll, 0.2% BSA, 100 mg/ml denatured salmon sperm DNA, 10% (wt/vol) dextran sulfate at 40°C, followed by one or more washes in 2X SSC, 25 mM Tris-HCl (pH 7.4), 5 mM EDTA, and 0.1% SDS at 50°C. Other conditions of low stringency that may be used are well known in the art (e.g., as employed for cross-species hybridizations). See, e.g., Ausubel *et al.* (eds.), 1993, CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, NY, and Kriegler, 1990, GENE TRANSFER AND EXPRESSION, A LABORATORY MANUAL, Stockton Press, NY; Shilo and Weinberg, 1981, *Proc Natl Acad Sci USA* 78: 6789-6792.

Conservative mutations

In addition to naturally-occurring allelic variants of the ORFX sequence that may exist in the population, the skilled artisan will further appreciate that changes can be introduced by mutation into the nucleotide sequence of any of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161), thereby leading to changes in the amino acid sequence of the encoded ORFX protein, without altering the functional ability of the ORFX protein. For example, nucleotide substitutions leading to amino acid substitutions at "non-essential" amino acid residues can be made in the sequence of any of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161). A "non-essential" amino acid residue is a residue that can be altered from the wild-type sequence of ORFX without altering the biological activity, whereas an "essential" amino acid residue is required for biological activity. For example, amino acid residues that are conserved among the ORFX proteins of the present invention, are predicted to be particularly unamenable to alteration.

Amino acid residues that are conserved among members of an ORFX family members are predicted to be less amenable to alteration. For example, an ORFX protein according to the present invention can contain at least one domain (*e.g.*, as shown in Table 1) that is a typically conserved region in an ORFX family member. As such, these conserved domains are not likely to be amenable to mutation. Other amino acid residues, however, (*e.g.*, those that are not conserved or only semi-conserved among members of the ORFX family) may not be as essential for activity and thus are more likely to be amenable to alteration.

Another aspect of the invention pertains to nucleic acid molecules encoding ORFX proteins that contain changes in amino acid residues that are not essential for activity. Such ORFX proteins differ in amino acid sequence from any of any of SEQ ID NO:2 n (wherein $n = 1$ to 3161), yet retain biological activity. In one embodiment, the isolated nucleic acid molecule comprises a nucleotide sequence encoding a protein, wherein the protein comprises an amino acid sequence at least about 75% homologous to the amino acid sequence of any of SEQ ID NO:2 n (wherein $n = 1$ to 3161). Preferably, the protein encoded by the nucleic acid is at least about 80% homologous to any of SEQ ID NO:2 n (wherein $n = 1$ to 3161), more preferably at least about 90%, 95%, 98%, and most preferably at least about 99% homologous to SEQ ID NO:2.

An isolated nucleic acid molecule encoding a ORFX protein homologous to the protein of any of SEQ ID NO:2 n (wherein $n = 1$ to 3161) can be created by introducing one or more nucleotide substitutions, additions or deletions into the corresponding nucleotide sequence, *i.e.* SEQ ID NO:2 n -1 for the corresponding n , such that one or more amino acid substitutions, additions or deletions are introduced into the encoded protein.

Mutations can be introduced into SEQ ID NO:2 n -1 (wherein $n = 1$ to 3161) by standard techniques, such as site-directed mutagenesis and PCR-mediated mutagenesis. Preferably, conservative amino acid substitutions are made at one or more predicted non-essential amino acid residues. A "conservative amino acid substitution" is one in which the amino acid residue is replaced with an amino acid residue having a similar side chain. Families of amino acid residues having similar side chains have been defined in the art. These families include amino acids with basic side chains (*e.g.*, lysine, arginine, histidine), acidic side chains (*e.g.*, aspartic acid, glutamic acid), uncharged polar side chains (*e.g.*, glycine, asparagine, glutamine, serine, threonine, tyrosine, cysteine), nonpolar side chains (*e.g.*, alanine, valine, leucine, isoleucine, proline,

phenylalanine, methionine, tryptophan), beta-branched side chains (*e.g.*, threonine, valine, isoleucine) and aromatic side chains (*e.g.*, tyrosine, phenylalanine, tryptophan, histidine). Thus, a predicted nonessential amino acid residue in ORFX is replaced with another amino acid residue from the same side chain family. Alternatively, in another embodiment, mutations can be introduced randomly along all or part of a ORFX coding sequence, such as by saturation mutagenesis, and the resultant mutants can be screened for ORFX biological activity to identify mutants that retain activity. Following mutagenesis of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161), the encoded protein can be expressed by any recombinant technology known in the art and the activity of the protein can be determined.

In one embodiment, a mutant ORFX protein can be assayed for (1) the ability to form protein:protein interactions with other ORFX proteins, other cell-surface proteins, or biologically active portions thereof, (2) complex formation between a mutant ORFX protein and a ORFX receptor; (3) the ability of a mutant ORFX protein to bind to an intracellular target protein or biologically active portion thereof; (*e.g.*, avidin proteins); (4) the ability to bind BRA protein; or (5) the ability to specifically bind an anti-ORFX protein antibody.

Antisense

Another aspect of the invention pertains to isolated antisense nucleic acid molecules that are hybridizable to or complementary to the nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161), or fragments, analogs or derivatives thereof. An "antisense" nucleic acid comprises a nucleotide sequence that is complementary to a "sense" nucleic acid encoding a protein, *e.g.*, complementary to the coding strand of a double-stranded cDNA molecule or complementary to an mRNA sequence. In specific aspects, antisense nucleic acid molecules are provided that comprise a sequence complementary to at least about 10, 25, 50, 100, 250 or 500 nucleotides or an entire ORFX coding strand, or to only a portion thereof. Nucleic acid molecules encoding fragments, homologs, derivatives and analogs of a ORFX protein of any of SEQ ID NO:2*n* (wherein *n* = 1 to 3161) or antisense nucleic acids complementary to a ORFX nucleic acid sequence of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161) are additionally provided.

In one embodiment, an antisense nucleic acid molecule is antisense to a "coding region" of the coding strand of a nucleotide sequence encoding ORFX. The term "coding region" refers to the region of the nucleotide sequence comprising codons which are translated into amino acid

residues (*e.g.*, the protein coding region of a human ORFX that corresponds to any of SEQ ID NO:2*n* (wherein *n* = 1 to 3161)). In another embodiment, the antisense nucleic acid molecule is antisense to a "noncoding region" of the coding strand of a nucleotide sequence encoding ORFX. The term "noncoding region" refers to 5' and 3' sequences which flank the coding region that are not translated into amino acids (*i.e.*, also referred to as 5' and 3' untranslated regions).

Given the coding strand sequences encoding ORFX disclosed herein (*e.g.*, SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161)), antisense nucleic acids of the invention can be designed according to the rules of Watson and Crick or Hoogsteen base pairing. The antisense nucleic acid molecule can be complementary to the entire coding region of ORFX mRNA, but more preferably is an oligonucleotide that is antisense to only a portion of the coding or noncoding region of ORFX mRNA. For example, the antisense oligonucleotide can be complementary to the region surrounding the translation start site of ORFX mRNA. An antisense oligonucleotide can be, for example, about 5, 10, 15, 20, 25, 30, 35, 40, 45 or 50 nucleotides in length. An antisense nucleic acid of the invention can be constructed using chemical synthesis or enzymatic ligation reactions using procedures known in the art. For example, an antisense nucleic acid (*e.g.*, an antisense oligonucleotide) can be chemically synthesized using naturally occurring nucleotides or variously modified nucleotides designed to increase the biological stability of the molecules or to increase the physical stability of the duplex formed between the antisense and sense nucleic acids, *e.g.*, phosphorothioate derivatives and acridine substituted nucleotides can be used.

Examples of modified nucleotides that can be used to generate the antisense nucleic acid include: 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xanthine, 4-acetylcytosine, 5-(carboxyhydroxymethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil,

3-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)w, and 2,6-diaminopurine. Alternatively, the antisense nucleic acid can be produced biologically using an expression vector into which a nucleic acid has been subcloned in an antisense orientation (*i.e.*, RNA transcribed from the inserted nucleic acid will be of an antisense orientation to a target nucleic acid of interest, described further in the following subsection).

The antisense nucleic acid molecules of the invention are typically administered to a subject or generated *in situ* such that they hybridize with or bind to cellular mRNA and/or genomic DNA encoding a ORFX protein to thereby inhibit expression of the protein, *e.g.*, by inhibiting transcription and/or translation. The hybridization can be by conventional nucleotide complementarity to form a stable duplex, or, for example, in the case of an antisense nucleic acid molecule that binds to DNA duplexes, through specific interactions in the major groove of the double helix. An example of a route of administration of antisense nucleic acid molecules of the invention includes direct injection at a tissue site. Alternatively, antisense nucleic acid molecules can be modified to target selected cells and then administered systemically. For example, for systemic administration, antisense molecules can be modified such that they specifically bind to receptors or antigens expressed on a selected cell surface, *e.g.*, by linking the antisense nucleic acid molecules to peptides or antibodies that bind to cell surface receptors or antigens. The antisense nucleic acid molecules can also be delivered to cells using the vectors described herein. To achieve sufficient intracellular concentrations of antisense molecules, vector constructs in which the antisense nucleic acid molecule is placed under the control of a strong pol II or pol III promoter are preferred.

In yet another embodiment, the antisense nucleic acid molecule of the invention is an α -anomeric nucleic acid molecule. An α -anomeric nucleic acid molecule forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual β -units, the strands run parallel to each other (Gaultier *et al.* (1987) *Nucleic Acids Res* 15: 6625-6641). The antisense nucleic acid molecule can also comprise a 2'-o-methylribonucleotide (Inoue *et al.* (1987) *Nucleic Acids Res* 15: 6131-6148) or a chimeric RNA -DNA analogue (Inoue *et al.* (1987) *FEBS Lett* 215: 327-330).

Ribozymes and PNA moieties

Such modifications include, by way of nonlimiting example, modified bases, and nucleic acids whose sugar phosphate backbones are modified or derivatized. These modifications are

carried out at least in part to enhance the chemical stability of the modified nucleic acid, such that they may be used, for example, as antisense binding nucleic acids in therapeutic applications in a subject.

In still another embodiment, an antisense nucleic acid of the invention is a ribozyme.

- 5 Ribozymes are catalytic RNA molecules with ribonuclease activity that are capable of cleaving a single-stranded nucleic acid, such as an mRNA, to which they have a complementary region. Thus, ribozymes (*e.g.*, hammerhead ribozymes (described in Haselhoff and Gerlach (1988) *Nature* 334:585-591)) can be used to catalytically cleave ORFX mRNA transcripts to thereby inhibit translation of ORFX mRNA. A ribozyme having specificity for a ORFX-encoding
- 10 nucleic acid can be designed based upon the nucleotide sequence of a ORFX DNA disclosed herein (*i.e.*, SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161)). For example, a derivative of a Tetrahymena L-19 IVS RNA can be constructed in which the nucleotide sequence of the active site is complementary to the nucleotide sequence to be cleaved in a ORFX-encoding mRNA. See, *e.g.*, Cech *et al.* U.S. Pat. No. 4,987,071; and Cech *et al.* U.S. Pat. No. 5,116,742.
- 15 Alternatively, ORFX mRNA can be used to select a catalytic RNA having a specific ribonuclease activity from a pool of RNA molecules. See, *e.g.*, Bartel *et al.*, (1993) *Science* 261:1411-1418.

- Alternatively, ORFX gene expression can be inhibited by targeting nucleotide sequences complementary to the regulatory region of the ORFX (*e.g.*, the ORFX promoter and/or
- 20 enhancers) to form triple helical structures that prevent transcription of the ORFX gene in target cells. See generally, Helene. (1991) *Anticancer Drug Des.* 6: 569-84; Helene. *et al.* (1992) *Ann. N.Y. Acad. Sci.* 660:27-36; and Maher (1992) *Bioassays* 14: 807-15.

- In various embodiments, the nucleic acids of ORFX can be modified at the base moiety, sugar moiety or phosphate backbone to improve, *e.g.*, the stability, hybridization, or solubility of
- 25 the molecule. For example, the deoxyribose phosphate backbone of the nucleic acids can be modified to generate peptide nucleic acids (see Hyrup *et al.* (1996) *Bioorg Med Chem* 4: 5-23). As used herein, the terms "peptide nucleic acids" or "PNAs" refer to nucleic acid mimics, *e.g.*, DNA mimics, in which the deoxyribose phosphate backbone is replaced by a pseudopeptide backbone and only the four natural nucleobases are retained. The neutral backbone of PNAs has
- 30 been shown to allow for specific hybridization to DNA and RNA under conditions of low ionic strength. The synthesis of PNA oligomers can be performed using standard solid phase peptide

synthesis protocols as described in Hyrup *et al.* (1996) above; Perry-O'Keefe *et al.* (1996) *PNAS* 93: 14670-675.

PNAs of ORFX can be used in therapeutic and diagnostic applications. For example, PNAs can be used as antisense or antigene agents for sequence-specific modulation of gene expression by, *e.g.*, inducing transcription or translation arrest or inhibiting replication. PNAs of ORFX can also be used, *e.g.*, in the analysis of single base pair mutations in a gene by, *e.g.*, PNA directed PCR clamping; as artificial restriction enzymes when used in combination with other enzymes, *e.g.*, S1 nucleases (Hyrup B. (1996) above); or as probes or primers for DNA sequence and hybridization (Hyrup *et al.* (1996), above; Perry-O'Keefe (1996), above).

In another embodiment, PNAs of ORFX can be modified, *e.g.*, to enhance their stability or cellular uptake, by attaching lipophilic or other helper groups to PNA, by the formation of PNA-DNA chimeras, or by the use of liposomes or other techniques of drug delivery known in the art. For example, PNA-DNA chimeras of ORFX can be generated that may combine the advantageous properties of PNA and DNA. Such chimeras allow DNA recognition enzymes, *e.g.*, RNase H and DNA polymerases, to interact with the DNA portion while the PNA portion would provide high binding affinity and specificity. PNA-DNA chimeras can be linked using linkers of appropriate lengths selected in terms of base stacking, number of bonds between the nucleobases, and orientation (Hyrup (1996) above). The synthesis of PNA-DNA chimeras can be performed as described in Hyrup (1996) above and Finn *et al.* (1996) *Nucl Acids Res* 24: 3357-63. For example, a DNA chain can be synthesized on a solid support using standard phosphoramidite coupling chemistry, and modified nucleoside analogs, *e.g.*, 5'-(4-methoxytrityl)amino-5'-deoxy-thymidine phosphoramidite, can be used between the PNA and the 5' end of DNA (Mag *et al.* (1989) *Nucl Acid Res* 17: 5973-88). PNA monomers are then coupled in a stepwise manner to produce a chimeric molecule with a 5' PNA segment and a 3' DNA segment (Finn *et al.* (1996) above). Alternatively, chimeric molecules can be synthesized with a 5' DNA segment and a 3' PNA segment. See, Petersen *et al.* (1975) *Bioorg Med Chem Lett* 5: 1119-1124.

In other embodiments, the oligonucleotide may include other appended groups such as peptides (*e.g.*, for targeting host cell receptors *in vivo*), or agents facilitating transport across the cell membrane (see, *e.g.*, Letsinger *et al.*, 1989, *Proc. Natl. Acad. Sci. U.S.A.* 86:6553-6556; Lemaitre *et al.*, 1987, *Proc. Natl. Acad. Sci.* 84:648-652; PCT Publication No. W088/09810) or

the blood-brain barrier (see, *e.g.*, PCT Publication No. W089/10134). In addition, oligonucleotides can be modified with hybridization triggered cleavage agents (See, *e.g.*, Krol *et al.*, 1988, *BioTechniques* 6:958-976) or intercalating agents. (See, *e.g.*, Zon, 1988, *Pharm. Res.* 5: 539-549). To this end, the oligonucleotide may be conjugated to another molecule, *e.g.*, a peptide, a hybridization triggered cross-linking agent, a transport agent, a hybridization-triggered cleavage agent, etc.

ORFX polypeptides

The novel protein of the invention includes the ORFX-like protein whose sequence is provided in any of SEQ ID NO:2n (wherein n = 1 to 3161). The invention also includes a mutant or variant protein any of whose residues may be changed from the corresponding residue shown in FIG. 1 while still encoding a protein that maintains its ORFX-like activities and physiological functions, or a functional fragment thereof. For example, the invention includes the polypeptides encoded by the variant ORFX nucleic acids described above. In the mutant or variant protein, up to 20% or more of the residues may be so changed.

In general, an ORFX-like variant that preserves ORFX-like function includes any variant in which residues at a particular position in the sequence have been substituted by other amino acids, and further include the possibility of inserting an additional residue or residues between two residues of the parent protein as well as the possibility of deleting one or more residues from the parent sequence. Any amino acid substitution, insertion, or deletion is encompassed by the invention. In favorable circumstances, the substitution is a conservative substitution as defined above. Furthermore, without limiting the scope of the invention, positions of any of SEQ ID NO:2n (wherein n = 1 to 3161) may be substituted such that a mutant or variant protein may include one or more substitutions

The invention also includes isolated ORFX proteins, and biologically active portions thereof, or derivatives, fragments, analogs or homologs thereof. Also provided are polypeptide fragments suitable for use as immunogens to raise anti-ORFX antibodies. In one embodiment, native ORFX proteins can be isolated from cells or tissue sources by an appropriate purification scheme using standard protein purification techniques. In another embodiment, ORFX proteins are produced by recombinant DNA techniques. Alternative to recombinant expression, a ORFX

protein or polypeptide can be synthesized chemically using standard peptide synthesis techniques.

An "isolated" or "purified" protein or biologically active portion thereof is substantially free of cellular material or other contaminating proteins from the cell or tissue source from which the ORFX protein is derived, or substantially free from chemical precursors or other chemicals when chemically synthesized. The language "substantially free of cellular material" includes preparations of ORFX protein in which the protein is separated from cellular components of the cells from which it is isolated or recombinantly produced. In one embodiment, the language "substantially free of cellular material" includes preparations of ORFX protein having less than about 30% (by dry weight) of non-ORFX protein (also referred to herein as a "contaminating protein"), more preferably less than about 20% of non-ORFX protein, still more preferably less than about 10% of non-ORFX protein, and most preferably less than about 5% non-ORFX protein. When the ORFX protein or biologically active portion thereof is recombinantly produced, it is also preferably substantially free of culture medium, *i.e.*, culture medium represents less than about 20%, more preferably less than about 10%, and most preferably less than about 5% of the volume of the protein preparation.

The language "substantially free of chemical precursors or other chemicals" includes preparations of ORFX protein in which the protein is separated from chemical precursors or other chemicals that are involved in the synthesis of the protein. In one embodiment, the language "substantially free of chemical precursors or other chemicals" includes preparations of ORFX protein having less than about 30% (by dry weight) of chemical precursors or non-ORFX chemicals, more preferably less than about 20% chemical precursors or non-ORFX chemicals, still more preferably less than about 10% chemical precursors or non-ORFX chemicals, and most preferably less than about 5% chemical precursors or non-ORFX chemicals.

Biologically active portions of a ORFX protein include peptides comprising amino acid sequences sufficiently homologous to or derived from the amino acid sequence of the ORFX protein, *e.g.*, the amino acid sequence shown in SEQ ID NO:2 that include fewer amino acids than the full length ORFX proteins, and exhibit at least one activity of a ORFX protein. Typically, biologically active portions comprise a domain or motif with at least one activity of the ORFX protein. A biologically active portion of a ORFX protein can be a polypeptide which is, for example, 10, 25, 50, 100 or more amino acids in length.

A biologically active portion of a ORFX protein of the present invention may contain at least one of the above-identified domains conserved between the FGF family of proteins. Moreover, other biologically active portions, in which other regions of the protein are deleted, can be prepared by recombinant techniques and evaluated for one or more of the functional activities of a native ORFX protein.

In an embodiment, the ORFX protein has an amino acid sequence shown in any of SEQ ID NO:2n (wherein n = 1 to 3161). In other embodiments, the ORFX protein is substantially homologous to any of SEQ ID NO:2n (wherein n = 1 to 3161) and retains the functional activity of the protein of any of SEQ ID NO:2n (wherein n = 1 to 3161), yet differs in amino acid sequence due to natural allelic variation or mutagenesis, as described in detail below. Accordingly, in another embodiment, the ORFX protein is a protein that comprises an amino acid sequence at least about 45% homologous, and more preferably about 55, 65, 70, 75, 80, 85, 90, 95, 98 or even 99% homologous to the amino acid sequence of any of SEQ ID NO:2n (wherein n = 1 to 3161) and retains the functional activity of the ORFX proteins of the corresponding polypeptide having the sequence of SEQ ID NO:2n (wherein n = 1 to 3161).

Determining homology between two or more sequences

To determine the percent homology of two amino acid sequences or of two nucleic acids, the sequences are aligned for optimal comparison purposes (e.g., gaps can be introduced in either of the sequences being compared for optimal alignment between the sequences). The amino acid residues or nucleotides at corresponding amino acid positions or nucleotide positions are then compared. When a position in the first sequence is occupied by the same amino acid residue or nucleotide as the corresponding position in the second sequence, then the molecules are homologous at that position (i.e., as used herein amino acid or nucleic acid "homology" is equivalent to amino acid or nucleic acid "identity").

The nucleic acid sequence homology may be determined as the degree of identity between two sequences. The homology may be determined using computer programs known in the art, such as GAP software provided in the GCG program package. See, *Needleman and Wunsch 1970 J Mol Biol* 48: 443-453. Using GCG GAP software with the following settings for nucleic acid sequence comparison: GAP creation penalty of 5.0 and GAP extension penalty of 0.3, the coding region of the analogous nucleic acid sequences referred to above exhibits a

degree of identity preferably of at least 70%, 75%, 80%, 85%, 90%, 95%, 98%, or 99%, with the CDS (encoding) part of the DNA sequence shown in SEQ ID NO:2n-1 (wherein n = 1 to 3161).

The term "sequence identity" refers to the degree to which two polynucleotide or polypeptide sequences are identical on a residue-by-residue basis over a particular region of comparison. The term "percentage of sequence identity" is calculated by comparing two optimally aligned sequences over that region of comparison, determining the number of positions at which the identical nucleic acid base (*e.g.*, A, T, C, G, U, or I, in the case of nucleic acids) occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the region of comparison (*i.e.*, the window size), and multiplying the result by 100 to yield the percentage of sequence identity. The term "substantial identity" as used herein denotes a characteristic of a polynucleotide sequence, wherein the polynucleotide comprises a sequence that has at least 80 percent sequence identity, preferably at least 85 percent identity and often 90 to 95 percent sequence identity, more usually at least 99 percent sequence identity as compared to a reference sequence over a comparison region. The term "percentage of positive residues" is calculated by comparing two optimally aligned sequences over that region of comparison, determining the number of positions at which the identical and conservative amino acid substitutions, as defined above, occur in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the region of comparison (*i.e.*, the window size), and multiplying the result by 100 to yield the percentage of positive residues.

Chimeric and fusion proteins

The invention also provides ORFX chimeric or fusion proteins. As used herein, a ORFX "chimeric protein" or "fusion protein" includes a ORFX polypeptide operatively linked to a non-ORFX polypeptide. A "ORFX polypeptide" refers to a polypeptide having an amino acid sequence corresponding to ORFX, whereas a "non-ORFX polypeptide" refers to a polypeptide having an amino acid sequence corresponding to a protein that is not substantially homologous to the ORFX protein, *e.g.*, a protein that is different from the ORFX protein and that is derived from the same or a different organism. Within a ORFX fusion protein the ORFX polypeptide can correspond to all or a portion of a ORFX protein. In one embodiment, a ORFX fusion protein comprises at least one biologically active portion of a ORFX protein. In another embodiment, a ORFX fusion protein comprises at least two biologically active portions of a

ORFX protein. Within the fusion protein, the term "operatively linked" is intended to indicate that the ORFX polypeptide and the non-ORFX polypeptide are fused in-frame to each other. The non-ORFX polypeptide can be fused to the N-terminus or C-terminus of the ORFX polypeptide.

5 For example, in one embodiment a ORFX fusion protein comprises a ORFX polypeptide operably linked to the extracellular domain of a second protein. Such fusion proteins can be further utilized in screening assays for compounds that modulate ORFX activity (such assays are described in detail below).

10 In another embodiment, the fusion protein is a GST-ORFX fusion protein in which the ORFX sequences are fused to the C-terminus of the GST (i.e., glutathione S-transferase) sequences. Such fusion proteins can facilitate the purification of recombinant ORFX.

15 In yet another embodiment, the fusion protein is a ORFX protein containing a heterologous signal sequence at its N-terminus. For example, the native ORFX signal sequence can be removed and replaced with a signal sequence from another protein. In certain host cells (e.g., mammalian host cells), expression and/or secretion of ORFX can be increased through use of a heterologous signal sequence.

20 In another embodiment, the fusion protein is a ORFX-immunoglobulin fusion protein in which the ORFX sequences comprising one or more domains are fused to sequences derived from a member of the immunoglobulin protein family. The ORFX-immunoglobulin fusion proteins of the invention can be incorporated into pharmaceutical compositions and administered to a subject to inhibit an interaction between a ORFX ligand and a ORFX protein on the surface of a cell, to thereby suppress ORFX-mediated signal transduction *in vivo*. In one nonlimiting example, a contemplated ORFX ligand of the invention is an ORFX receptor. The ORFX-immunoglobulin fusion proteins can be used to modulate the bioavailability of a ORFX
25 cognate ligand. Inhibition of the ORFX ligand/ORFX interaction may be useful therapeutically for both the treatment of proliferative and differentiative disorders, as well as modulating (e.g., promoting or inhibiting) cell survival. Moreover, the ORFX-immunoglobulin fusion proteins of the invention can be used as immunogens to produce anti-ORFX antibodies in a subject, to purify ORFX ligands, and in screening assays to identify molecules that inhibit the interaction of
30 ORFX with a ORFX ligand.

A ORFX chimeric or fusion protein of the invention can be produced by standard recombinant DNA techniques. For example, DNA fragments coding for the different polypeptide sequences are ligated together in-frame in accordance with conventional techniques, *e.g.*, by employing blunt-ended or stagger-ended termini for ligation, restriction enzyme digestion to provide for appropriate termini, filling-in of cohesive ends as appropriate, alkaline phosphatase treatment to avoid undesirable joining, and enzymatic ligation. In another embodiment, the fusion gene can be synthesized by conventional techniques including automated DNA synthesizers. Alternatively, PCR amplification of gene fragments can be carried out using anchor primers that give rise to complementary overhangs between two consecutive gene fragments that can subsequently be annealed and reamplified to generate a chimeric gene sequence (see, for example, Ausubel et al. (eds.) CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, 1992). Moreover, many expression vectors are commercially available that already encode a fusion moiety (*e.g.*, a GST polypeptide). A ORFX-encoding nucleic acid can be cloned into such an expression vector such that the fusion moiety is linked in-frame to the ORFX protein.

ORFX agonists and antagonists

The present invention also pertains to variants of the ORFX proteins that function as either ORFX agonists (mimetics) or as ORFX antagonists. Variants of the ORFX protein can be generated by mutagenesis, *e.g.*, discrete point mutation or truncation of the ORFX protein. An agonist of the ORFX protein can retain substantially the same, or a subset of, the biological activities of the naturally occurring form of the ORFX protein. An antagonist of the ORFX protein can inhibit one or more of the activities of the naturally occurring form of the ORFX protein by, for example, competitively binding to a downstream or upstream member of a cellular signaling cascade which includes the ORFX protein. Thus, specific biological effects can be elicited by treatment with a variant of limited function. In one embodiment, treatment of a subject with a variant having a subset of the biological activities of the naturally occurring form of the protein has fewer side effects in a subject relative to treatment with the naturally occurring form of the ORFX proteins.

Variants of the ORFX protein that function as either ORFX agonists (mimetics) or as ORFX antagonists can be identified by screening combinatorial libraries of mutants, *e.g.*, truncation mutants, of the ORFX protein for ORFX protein agonist or antagonist activity. In one

embodiment, a variegated library of ORFX variants is generated by combinatorial mutagenesis at the nucleic acid level and is encoded by a variegated gene library. A variegated library of ORFX variants can be produced by, for example, enzymatically ligating a mixture of synthetic oligonucleotides into gene sequences such that a degenerate set of potential ORFX sequences is expressible as individual polypeptides, or alternatively, as a set of larger fusion proteins (e.g., for phage display) containing the set of ORFX sequences therein. There are a variety of methods which can be used to produce libraries of potential ORFX variants from a degenerate oligonucleotide sequence. Chemical synthesis of a degenerate gene sequence can be performed in an automatic DNA synthesizer, and the synthetic gene then ligated into an appropriate expression vector. Use of a degenerate set of genes allows for the provision, in one mixture, of all of the sequences encoding the desired set of potential ORFX sequences. Methods for synthesizing degenerate oligonucleotides are known in the art (see, e.g., Narang (1983) *Tetrahedron* 39:3; Itakura *et al.* (1984) *Annu Rev Biochem* 53:323; Itakura *et al.* (1984) *Science* 198:1056; Ike *et al.* (1983) *Nucl Acid Res* 11:477.

Polypeptide libraries

In addition, libraries of fragments of the ORFX protein coding sequence can be used to generate a variegated population of ORFX fragments for screening and subsequent selection of variants of a ORFX protein. In one embodiment, a library of coding sequence fragments can be generated by treating a double stranded PCR fragment of a ORFX coding sequence with a nuclease under conditions wherein nicking occurs only about once per molecule, denaturing the double stranded DNA, renaturing the DNA to form double stranded DNA that can include sense/antisense pairs from different nicked products, removing single stranded portions from reformed duplexes by treatment with S1 nuclease, and ligating the resulting fragment library into an expression vector. By this method, an expression library can be derived which encodes N-terminal and internal fragments of various sizes of the ORFX protein.

Several techniques are known in the art for screening gene products of combinatorial libraries made by point mutations or truncation, and for screening cDNA libraries for gene products having a selected property. Such techniques are adaptable for rapid screening of the gene libraries generated by the combinatorial mutagenesis of ORFX proteins. The most widely used techniques, which are amenable to high throughput analysis, for screening large gene libraries typically include cloning the gene library into replicable expression vectors,

transforming appropriate cells with the resulting library of vectors, and expressing the combinatorial genes under conditions in which detection of a desired activity facilitates isolation of the vector encoding the gene whose product was detected. Recursive ensemble mutagenesis (REM), a new technique that enhances the frequency of functional mutants in the libraries, can be used in combination with the screening assays to identify ORFX variants (Arkin and Yourvan (1992) PNAS 89:7811-7815; Delgrave *et al.* (1993) Protein Engineering 6:327-331).

Anti-ORFX Antibodies

The invention further encompasses antibodies and antibody fragments, such as F_{ab} or $(F_{ab})_2$, that bind immunospecifically to any of the proteins of the invention.

An isolated ORFX protein, or a portion or fragment thereof, can be used as an immunogen to generate antibodies that bind ORFX using standard techniques for polyclonal and monoclonal antibody preparation. Full-length ORFX protein can be used. Alternatively, the invention provides antigenic peptide fragments of ORFX for use as immunogens. The antigenic peptide of ORFX comprises at least 4 amino acid residues of the amino acid sequence shown in any of SEQ ID NO:2n (wherein n = 1 to 3161). The antigenic peptide encompasses an epitope of ORFX such that an antibody raised against the peptide forms a specific immune complex with ORFX. The antigenic peptide may comprise at least 6 aa residues, at least 8 aa residues, at least 10 aa residues, at least 15 aa residues, at least 20 aa residues, or at least 30 aa residues. In one embodiment of the invention, the antigenic peptide comprises a polypeptide comprising at least 6 contiguous amino acids of any of SEQ ID NO:2n (wherein n = 1 to 3161).

In an embodiment of the invention, epitopes encompassed by the antigenic peptide are regions of ORFX that are located on the surface of the protein, *e.g.*, hydrophilic regions. As a means for targeting antibody production, hydropathy plots showing regions of hydrophilicity and hydrophobicity may be generated by any method well known in the art, including, for example, the Kyte Doolittle or the Hopp Woods methods, either with or without Fourier transformation. See, *e.g.*, Hopp and Woods, 1981, Proc. Nat. Acad. Sci. USA 78: 3824-3828; Kyte and Doolittle 1982, J. Mol. Biol. 157: 105-142, each incorporated herein by reference in their entirety.

As disclosed herein, an ORFX protein sequence of any of SEQ ID NO:2n (wherein n = 1 to 3161), or derivatives, fragments, analogs or homologs thereof, may be utilized as immunogens in the generation of antibodies that immunospecifically-bind these protein components. The term "antibody" as used herein refers to immunoglobulin molecules and

immunologically active portions of immunoglobulin molecules, i.e., molecules that contain an antigen binding site that specifically binds (immunoreacts with) an antigen, such as ORFX. Such antibodies include, but are not limited to, polyclonal, monoclonal, chimeric, single chain, F_{ab} and $F_{(ab)_2}$ fragments, and an F_{ab} expression library. In a specific embodiment, antibodies to human ORFX proteins are disclosed. Various procedures known within the art may be used for the production of polyclonal or monoclonal antibodies to a ORFX protein sequence of any of SEQ ID NO:2n (wherein n = 1 to 3161) or derivative, fragment, analog or homolog thereof. Some of these proteins are discussed below.

For the production of polyclonal antibodies, various suitable host animals (e.g., rabbit, goat, mouse or other mammal) may be immunized by injection with the native protein, or a synthetic variant thereof, or a derivative of the foregoing. An appropriate immunogenic preparation can contain, for example, recombinantly expressed ORFX protein or a chemically synthesized ORFX polypeptide. The preparation can further include an adjuvant. Various adjuvants used to increase the immunological response include, but are not limited to, Freund's (complete and incomplete), mineral gels (e.g., aluminum hydroxide), surface active substances (e.g., lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, dinitrophenol, etc.), human adjuvants such as *Bacille Calmette-Guerin* and *Corynebacterium parvum*, or similar immunostimulatory agents. If desired, the antibody molecules directed against ORFX can be isolated from the mammal (e.g., from the blood) and further purified by well known techniques, such as protein A chromatography to obtain the IgG fraction.

The term "monoclonal antibody" or "monoclonal antibody composition", as used herein, refers to a population of antibody molecules that contain only one species of an antigen binding site capable of immunoreacting with a particular epitope of ORFX. A monoclonal antibody composition thus typically displays a single binding affinity for a particular ORFX protein with which it immunoreacts. For preparation of monoclonal antibodies directed towards a particular ORFX protein, or derivatives, fragments, analogs or homologs thereof, any technique that provides for the production of antibody molecules by continuous cell line culture may be utilized. Such techniques include, but are not limited to, the hybridoma technique (see Kohler & Milstein, 1975 *Nature* 256: 495-497); the trioma technique; the human B-cell hybridoma technique (see Kozbor, *et al.*, 1983 *Immunol Today* 4: 72) and the EBV hybridoma technique to produce human monoclonal antibodies (see Cole, *et al.*, 1985 In: MONOCLONAL ANTIBODIES AND

CANCER THERAPY, Alan R. Liss, Inc., pp. 77-96). Human monoclonal antibodies may be utilized in the practice of the present invention and may be produced by using human hybridomas (see Cote, *et al.*, 1983. *Proc Natl Acad Sci USA* 80: 2026-2030) or by transforming human B-cells with Epstein Barr Virus *in vitro* (see Cole, *et al.*, 1985 In: MONOCLONAL ANTIBODIES AND
5 CANCER THERAPY, Alan R. Liss, Inc., pp. 77-96). Each of the above citations are incorporated herein by reference in their entirety

According to the invention, techniques can be adapted for the production of single-chain antibodies specific to a ORFX protein (see *e.g.*, U.S. Patent No. 4,946,778). In addition, methods can be adapted for the construction of F_{ab} expression libraries (see *e.g.*, Huse, *et al.*,
10 1989 *Science* 246: 1275-1281) to allow rapid and effective identification of monoclonal F_{ab} fragments with the desired specificity for a ORFX protein or derivatives, fragments, analogs or homologs thereof. Non-human antibodies can be "humanized" by techniques well known in the art. See *e.g.*, U.S. Patent No. 5,225,539. Each of the above citations are incorporated herein by reference. Antibody fragments that contain the idiotypes to a ORFX protein may be produced by
15 techniques known in the art including, but not limited to: (i) an F_(ab)2 fragment produced by pepsin digestion of an antibody molecule; (ii) an F_{ab} fragment generated by reducing the disulfide bridges of an F_(ab)2 fragment; (iii) an F_{ab} fragment generated by the treatment of the antibody molecule with papain and a reducing agent and (iv) F_v fragments.

Additionally, recombinant anti-ORFX antibodies, such as chimeric and humanized
20 monoclonal antibodies, comprising both human and non-human portions, which can be made using standard recombinant DNA techniques, are within the scope of the invention. Such chimeric and humanized monoclonal antibodies can be produced by recombinant DNA techniques known in the art, for example using methods described in PCT International Application No. PCT/US86/02269; European Patent Application No. 184,187; European Patent
25 Application No. 171,496; European Patent Application No. 173,494; PCT International Publication No. WO 86/01533; U.S. Pat. No. 4,816,567; European Patent Application No. 125,023; Better *et al.* (1988) *Science* 240:1041-1043; Liu *et al.* (1987) *PNAS* 84:3439-3443; Liu *et al.* (1987) *J Immunol.* 139:3521-3526; Sun *et al.* (1987) *PNAS* 84:214-218; Nishimura *et al.* (1987) *Cancer Res* 47:999-1005; Wood *et al.* (1985) *Nature* 314:446-449; Shaw *et al.* (1988), *J.*
30 *Natl Cancer Inst* 80:1553-1559; Morrison (1985) *Science* 229:1202-1207; Oi *et al.* (1986) *BioTechniques* 4:214; U.S. Pat. No. 5,225,539; Jones *et al.* (1986) *Nature* 321:552-525;

Verhoeyan *et al.* (1988) *Science* 239:1534; and Beidler *et al.* (1988) *J Immunol* 141:4053-4060. Each of the above citations are incorporated herein by reference.

In one embodiment, methods for the screening of antibodies that possess the desired specificity include, but are not limited to, enzyme-linked immunosorbent assay (ELISA) and other immunologically-mediated techniques known within the art. In a specific embodiment, selection of antibodies that are specific to a particular domain of a ORFX protein is facilitated by generation of hybridomas that bind to the fragment of a ORFX protein possessing such a domain. Antibodies that are specific for one or more domains within a ORFX protein, *e.g.*, the domain spanning the first fifty amino-terminal residues specific to ORFX when compared to FGF-9, or derivatives, fragments, analogs or homologs thereof, are also provided herein.

Anti-ORFX antibodies may be used in methods known within the art relating to the localization and/or quantitation of a ORFX protein (*e.g.*, for use in measuring levels of the ORFX protein within appropriate physiological samples, for use in diagnostic methods, for use in imaging the protein, and the like). In a given embodiment, antibodies for ORFX proteins, or derivatives, fragments, analogs or homologs thereof, that contain the antibody derived binding domain, are utilized as pharmacologically-active compounds [hereinafter "Therapeutics"].

An anti-ORFX antibody (*e.g.*, monoclonal antibody) can be used to isolate ORFX by standard techniques, such as affinity chromatography or immunoprecipitation. An anti-ORFX antibody can facilitate the purification of natural ORFX from cells and of recombinantly produced ORFX expressed in host cells. Moreover, an anti-ORFX antibody can be used to detect ORFX protein (*e.g.*, in a cellular lysate or cell supernatant) in order to evaluate the abundance and pattern of expression of the ORFX protein. Anti-ORFX antibodies can be used diagnostically to monitor protein levels in tissue as part of a clinical testing procedure, *e.g.*, to, for example, determine the efficacy of a given treatment regimen. Detection can be facilitated by coupling (*i.e.*, physically linking) the antibody to a detectable substance. Examples of detectable substances include various enzymes, prosthetic groups, fluorescent materials, luminescent materials, bioluminescent materials, and radioactive materials. Examples of suitable enzymes include horseradish peroxidase, alkaline phosphatase, β -galactosidase, or acetylcholinesterase; examples of suitable prosthetic group complexes include streptavidin/biotin and avidin/biotin; examples of suitable fluorescent materials include umbelliferone, fluorescein, fluorescein isothiocyanate, rhodamine, dichlorotriazinylamine fluorescein, dansyl chloride or

phycoerythrin; an example of a luminescent material includes luminol; examples of bioluminescent materials include luciferase, luciferin, and aequorin, and examples of suitable radioactive material include ^{125}I , ^{131}I , ^{35}S or ^3H .

ORFX Recombinant Vectors and Host Cells

5 Another aspect of the invention pertains to vectors, preferably expression vectors, containing a nucleic acid encoding ORFX protein, or derivatives, fragments, analogs or homologs thereof. As used herein, the term "vector" refers to a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked. One type of vector is a "plasmid", which refers to a circular double stranded DNA loop into which additional DNA segments can
10 be ligated. Another type of vector is a viral vector, wherein additional DNA segments can be ligated into the viral genome. Certain vectors are capable of autonomous replication in a host cell into which they are introduced (*e.g.*, bacterial vectors having a bacterial origin of replication and episomal mammalian vectors). Other vectors (*e.g.*, non-episomal mammalian vectors) are integrated into the genome of a host cell upon introduction into the host cell, and thereby are
15 replicated along with the host genome. Moreover, certain vectors are capable of directing the expression of genes to which they are operatively linked. Such vectors are referred to herein as "expression vectors". In general, expression vectors of utility in recombinant DNA techniques are often in the form of plasmids. In the present specification, "plasmid" and "vector" can be used interchangeably as the plasmid is the most commonly used form of vector. However, the
20 invention is intended to include such other forms of expression vectors, such as viral vectors (*e.g.*, replication defective retroviruses, adenoviruses and adeno-associated viruses), which serve equivalent functions.

The recombinant expression vectors of the invention comprise a nucleic acid of the invention in a form suitable for expression of the nucleic acid in a host cell, which means that the
25 recombinant expression vectors include one or more regulatory sequences, selected on the basis of the host cells to be used for expression, that is operatively linked to the nucleic acid sequence to be expressed. Within a recombinant expression vector, "operably linked" is intended to mean that the nucleotide sequence of interest is linked to the regulatory sequence(s) in a manner that allows for expression of the nucleotide sequence (*e.g.*, in an *in vitro* transcription/translation
30 system or in a host cell when the vector is introduced into the host cell). The term "regulatory sequence" is intended to include promoters, enhancers and other expression control elements

(*e.g.*, polyadenylation signals). Such regulatory sequences are described, for example, in Goeddel; GENE EXPRESSION TECHNOLOGY: METHODS IN ENZYMOLOGY 185, Academic Press, San Diego, Calif. (1990). Regulatory sequences include those that direct constitutive expression of a nucleotide sequence in many types of host cell and those that direct expression of the nucleotide sequence only in certain host cells (*e.g.*, tissue-specific regulatory sequences). It will be appreciated by those skilled in the art that the design of the expression vector can depend on such factors as the choice of the host cell to be transformed, the level of expression of protein desired, etc. The expression vectors of the invention can be introduced into host cells to thereby produce proteins or peptides, including fusion proteins or peptides, encoded by nucleic acids as described herein (*e.g.*, ORFX proteins, mutant forms of ORFX, fusion proteins, etc.).

The recombinant expression vectors of the invention can be designed for expression of ORFX in prokaryotic or eukaryotic cells. For example, ORFX can be expressed in bacterial cells such as *E. coli*, insect cells (using baculovirus expression vectors) yeast cells or mammalian cells. Suitable host cells are discussed further in Goeddel, GENE EXPRESSION TECHNOLOGY: METHODS IN ENZYMOLOGY 185, Academic Press, San Diego, Calif. (1990). Alternatively, the recombinant expression vector can be transcribed and translated *in vitro*, for example using T7 promoter regulatory sequences and T7 polymerase.

Expression of proteins in prokaryotes is most often carried out in *E. coli* with vectors containing constitutive or inducible promoters directing the expression of either fusion or non-fusion proteins. Fusion vectors add a number of amino acids to a protein encoded therein, usually to the amino terminus of the recombinant protein. Such fusion vectors typically serve three purposes: (1) to increase expression of recombinant protein; (2) to increase the solubility of the recombinant protein; and (3) to aid in the purification of the recombinant protein by acting as a ligand in affinity purification. Often, in fusion expression vectors, a proteolytic cleavage site is introduced at the junction of the fusion moiety and the recombinant protein to enable separation of the recombinant protein from the fusion moiety subsequent to purification of the fusion protein. Such enzymes, and their cognate recognition sequences, include Factor Xa, thrombin and enterokinase. Typical fusion expression vectors include pGEX (Pharmacia Biotech Inc; Smith and Johnson (1988) *Gene* 67:31-40), pMAL (New England Biolabs, Beverly, Mass.) and pRIT5 (Pharmacia, Piscataway, N.J.) that fuse glutathione S-transferase (GST), maltose E binding protein, or protein A, respectively, to the target recombinant protein.

Examples of suitable inducible non-fusion *E. coli* expression vectors include pTrc (Amrann *et al.*, (1988) *Gene* 69:301-315) and pET 11d (Studier *et al.*, GENE EXPRESSION TECHNOLOGY: METHODS IN ENZYMOLOGY 185, Academic Press, San Diego, Calif. (1990) 60-89).

5 One strategy to maximize recombinant protein expression in *E. coli* is to express the protein in a host bacteria with an impaired capacity to proteolytically cleave the recombinant protein. See, Gottesman, GENE EXPRESSION TECHNOLOGY: METHODS IN ENZYMOLOGY 185, Academic Press, San Diego, Calif. (1990) 119-128. Another strategy is to alter the nucleic acid sequence of the nucleic acid to be inserted into an expression vector so that the individual codons
10 for each amino acid are those preferentially utilized in *E. coli* (Wada *et al.*, (1992) *Nucleic Acids Res.* 20:2111-2118). Such alteration of nucleic acid sequences of the invention can be carried out by standard DNA synthesis techniques.

In another embodiment, the ORFX expression vector is a yeast expression vector. Examples of vectors for expression in yeast *S. cerevisiae* include pYepSec1 (Baldari, *et al.*,
15 (1987) *EMBO J* 6:229-234), pMFa (Kurjan and Herskowitz, (1982) *Cell* 30:933-943), pJRY88 (Schultz *et al.*, (1987) *Gene* 54:113-123), pYES2 (Invitrogen Corporation, San Diego, Calif.), and picZ (InVitrogen Corp, San Diego, Calif.).

Alternatively, ORFX can be expressed in insect cells using baculovirus expression vectors. Baculovirus vectors available for expression of proteins in cultured insect cells (*e.g.*,
20 SF9 cells) include the pAc series (Smith *et al.* (1983) *Mol Cell Biol* 3:2156-2165) and the pVL series (Lucklow and Summers (1989) *Virology* 170:31-39).

In yet another embodiment, a nucleic acid of the invention is expressed in mammalian cells using a mammalian expression vector. Examples of mammalian expression vectors include pCDM8 (Seed (1987) *Nature* 329:840) and pMT2PC (Kaufman *et al.* (1987) *EMBO J*
25 6: 187-195). When used in mammalian cells, the expression vector's control functions are often provided by viral regulatory elements. For example, commonly used promoters are derived from polyoma, Adenovirus 2, cytomegalovirus and Simian Virus 40. For other suitable expression systems for both prokaryotic and eukaryotic cells. See, *e.g.*, Chapters 16 and 17 of Sambrook *et al.*, MOLECULAR CLONING: A LABORATORY MANUAL. 2nd ed., Cold Spring Harbor Laboratory,
30 Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1989.

In another embodiment, the recombinant mammalian expression vector is capable of directing expression of the nucleic acid preferentially in a particular cell type (*e.g.*, tissue-specific regulatory elements are used to express the nucleic acid). Tissue-specific regulatory elements are known in the art. Non-limiting examples of suitable tissue-specific promoters include the albumin promoter (liver-specific; Pinkert *et al.* (1987) *Genes Dev* 1:268-277), lymphoid-specific promoters (Calame and Eaton (1988) *Adv Immunol* 43:235-275), in particular promoters of T cell receptors (Winoto and Baltimore (1989) *EMBO J* 8:729-733) and immunoglobulins (Banerji *et al.* (1983) *Cell* 33:729-740; Queen and Baltimore (1983) *Cell* 33:741-748), neuron-specific promoters (*e.g.*, the neurofilament promoter; Byrne and Ruddle (1989) *PNAS* 86:5473-5477), pancreas-specific promoters (Edlund *et al.* (1985) *Science* 230:912-916), and mammary gland-specific promoters (*e.g.*, milk whey promoter; U.S. Pat. No. 4,873,316 and European Application Publication No. 264,166). Developmentally-regulated promoters are also encompassed, *e.g.*, the murine hox promoters (Kessel and Gruss (1990) *Science* 249:374-379) and the α -fetoprotein promoter (Campes and Tilghman (1989) *Genes Dev* 3:537-546).

The invention further provides a recombinant expression vector comprising a DNA molecule of the invention cloned into the expression vector in an antisense orientation. That is, the DNA molecule is operatively linked to a regulatory sequence in a manner that allows for expression (by transcription of the DNA molecule) of an RNA molecule that is antisense to ORFX mRNA. Regulatory sequences operatively linked to a nucleic acid cloned in the antisense orientation can be chosen that direct the continuous expression of the antisense RNA molecule in a variety of cell types, for instance viral promoters and/or enhancers, or regulatory sequences can be chosen that direct constitutive, tissue specific or cell type specific expression of antisense RNA. The antisense expression vector can be in the form of a recombinant plasmid, phagemid or attenuated virus in which antisense nucleic acids are produced under the control of a high efficiency regulatory region, the activity of which can be determined by the cell type into which the vector is introduced. For a discussion of the regulation of gene expression using antisense genes see Weintraub *et al.*, "Antisense RNA as a molecular tool for genetic analysis," Reviews--Trends in Genetics, Vol. 1(1) 1986.

Another aspect of the invention pertains to host cells into which a recombinant expression vector of the invention has been introduced. The terms "host cell" and "recombinant

host cell" are used interchangeably herein. It is understood that such terms refer not only to the particular subject cell but to the progeny or potential progeny of such a cell. Because certain modifications may occur in succeeding generations due to either mutation or environmental influences, such progeny may not, in fact, be identical to the parent cell, but are still included within the scope of the term as used herein.

A host cell can be any prokaryotic or eukaryotic cell. For example, ORFX protein can be expressed in bacterial cells such as *E. coli*, insect cells, yeast or mammalian cells (such as Chinese hamster ovary cells (CHO) or COS cells). Other suitable host cells are known to those skilled in the art.

Vector DNA can be introduced into prokaryotic or eukaryotic cells via conventional transformation or transfection techniques. As used herein, the terms "transformation" and "transfection" are intended to refer to a variety of art-recognized techniques for introducing foreign nucleic acid (*e.g.*, DNA) into a host cell, including calcium phosphate or calcium chloride co-precipitation, DEAE-dextran-mediated transfection, lipofection, or electroporation. Suitable methods for transforming or transfecting host cells can be found in Sambrook, *et al.* (MOLECULAR CLONING: A LABORATORY MANUAL. 2nd ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1989), and other laboratory manuals.

For stable transfection of mammalian cells, it is known that, depending upon the expression vector and transfection technique used, only a small fraction of cells may integrate the foreign DNA into their genome. In order to identify and select these integrants, a gene that encodes a selectable marker (*e.g.*, resistance to antibiotics) is generally introduced into the host cells along with the gene of interest. Various selectable markers include those that confer resistance to drugs, such as G418, hygromycin and methotrexate. Nucleic acid encoding a selectable marker can be introduced into a host cell on the same vector as that encoding ORFX or can be introduced on a separate vector. Cells stably transfected with the introduced nucleic acid can be identified by drug selection (*e.g.*, cells that have incorporated the selectable marker gene will survive, while the other cells die).

A host cell of the invention, such as a prokaryotic or eukaryotic host cell in culture, can be used to produce (*i.e.*, express) ORFX protein. Accordingly, the invention further provides methods for producing ORFX protein using the host cells of the invention. In one embodiment,

the method comprises culturing the host cell of invention (into which a recombinant expression vector encoding ORFX has been introduced) in a suitable medium such that ORFX protein is produced. In another embodiment, the method further comprises isolating ORFX from the medium or the host cell.

5 **Transgenic animals**

The host cells of the invention can also be used to produce nonhuman transgenic animals. For example, in one embodiment, a host cell of the invention is a fertilized oocyte or an embryonic stem cell into which ORFX-coding sequences have been introduced. Such host cells can then be used to create non-human transgenic animals in which exogenous ORFX sequences have been introduced into their genome or homologous recombinant animals in which endogenous ORFX sequences have been altered. Such animals are useful for studying the function and/or activity of ORFX and for identifying and/or evaluating modulators of ORFX activity. As used herein, a "transgenic animal" is a non-human animal, preferably a mammal, more preferably a rodent such as a rat or mouse, in which one or more of the cells of the animal includes a transgene. Other examples of transgenic animals include non-human primates, sheep, dogs, cows, goats, chickens, amphibians, etc. A transgene is exogenous DNA that is integrated into the genome of a cell from which a transgenic animal develops and that remains in the genome of the mature animal, thereby directing the expression of an encoded gene product in one or more cell types or tissues of the transgenic animal. As used herein, a "homologous recombinant animal" is a non-human animal, preferably a mammal, more preferably a mouse, in which an endogenous ORFX gene has been altered by homologous recombination between the endogenous gene and an exogenous DNA molecule introduced into a cell of the animal, *e.g.*, an embryonic cell of the animal, prior to development of the animal.

A transgenic animal of the invention can be created by introducing ORFX-encoding nucleic acid into the male pronuclei of a fertilized oocyte, *e.g.*, by microinjection, retroviral infection, and allowing the oocyte to develop in a pseudopregnant female foster animal. The human ORFX DNA sequence of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161) can be introduced as a transgene into the genome of a non-human animal. Alternatively, a nonhuman homologue of the human ORFX gene, such as a mouse ORFX gene, can be isolated based on hybridization to the human ORFX cDNA (described further above) and used as a transgene. Intronic sequences and polyadenylation signals can also be included in the transgene to increase the efficiency of

expression of the transgene. A tissue-specific regulatory sequence(s) can be operably linked to the ORFX transgene to direct expression of ORFX protein to particular cells. Methods for generating transgenic animals via embryo manipulation and microinjection, particularly animals such as mice, have become conventional in the art and are described, for example, in U.S. Pat. Nos. 4,736,866; 4,870,009; and 4,873,191; and Hogan 1986, In: MANIPULATING THE MOUSE EMBRYO, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. Similar methods are used for production of other transgenic animals. A transgenic founder animal can be identified based upon the presence of the ORFX transgene in its genome and/or expression of ORFX mRNA in tissues or cells of the animals. A transgenic founder animal can then be used to breed additional animals carrying the transgene. Moreover, transgenic animals carrying a transgene encoding ORFX can further be bred to other transgenic animals carrying other transgenes.

To create a homologous recombinant animal, a vector is prepared which contains at least a portion of a ORFX gene into which a deletion, addition or substitution has been introduced to thereby alter, *e.g.*, functionally disrupt, the ORFX gene. The ORFX gene can be a human gene (*e.g.*, SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161)), but more preferably, is a non-human homologue of a human ORFX gene. For example, a mouse homologue of human ORFX gene of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161) can be used to construct a homologous recombination vector suitable for altering an endogenous ORFX gene in the mouse genome. In one embodiment, the vector is designed such that, upon homologous recombination, the endogenous ORFX gene is functionally disrupted (*i.e.*, no longer encodes a functional protein; also referred to as a "knock out" vector).

Alternatively, the vector can be designed such that, upon homologous recombination, the endogenous ORFX gene is mutated or otherwise altered but still encodes functional protein (*e.g.*, the upstream regulatory region can be altered to thereby alter the expression of the endogenous ORFX protein). In the homologous recombination vector, the altered portion of the ORFX gene is flanked at its 5' and 3' ends by additional nucleic acid of the ORFX gene to allow for homologous recombination to occur between the exogenous ORFX gene carried by the vector and an endogenous ORFX gene in an embryonic stem cell. The additional flanking ORFX nucleic acid is of sufficient length for successful homologous recombination with the endogenous gene. Typically, several kilobases of flanking DNA (both at the 5' and 3' ends) are included in the vector. See *e.g.*, Thomas *et al.* (1987) *Cell* 51:503 for a description of

homologous recombination vectors. The vector is introduced into an embryonic stem cell line (e.g., by electroporation) and cells in which the introduced ORFX gene has homologously recombined with the endogenous ORFX gene are selected (see e.g., Li *et al.* (1992) *Cell* 69:915).

The selected cells are then injected into a blastocyst of an animal (e.g., a mouse) to form aggregation chimeras. See e.g., Bradley 1987, In: TERATOCARCINOMAS AND EMBRYONIC STEM CELLS: A PRACTICAL APPROACH, Robertson, ed. IRL, Oxford, pp. 113-152. A chimeric embryo can then be implanted into a suitable pseudopregnant female foster animal and the embryo brought to term. Progeny harboring the homologously recombined DNA in their germ cells can be used to breed animals in which all cells of the animal contain the homologously recombined DNA by germline transmission of the transgene. Methods for constructing homologous recombination vectors and homologous recombinant animals are described further in Bradley (1991) *Curr Opin Biotechnol* 2:823-829; PCT International Publication Nos.: WO 90/11354; WO 91/01140; WO 92/0968; and WO 93/04169.

In another embodiment, transgenic non-humans animals can be produced that contain selected systems that allow for regulated expression of the transgene. One example of such a system is the cre/loxP recombinase system of bacteriophage P1. For a description of the cre/loxP recombinase system, see, e.g., Lakso *et al.* (1992) *PNAS* 89:6232-6236. Another example of a recombinase system is the FLP recombinase system of *Saccharomyces cerevisiae* (O'Gorman *et al.* (1991) *Science* 251:1351-1355. If a cre/loxP recombinase system is used to regulate expression of the transgene, animals containing transgenes encoding both the Cre recombinase and a selected protein are required. Such animals can be provided through the construction of "double" transgenic animals, e.g., by mating two transgenic animals, one containing a transgene encoding a selected protein and the other containing a transgene encoding a recombinase.

Clones of the non-human transgenic animals described herein can also be produced according to the methods described in Wilmot *et al.* (1997) *Nature* 385:810-813. In brief, a cell, e.g., a somatic cell, from the transgenic animal can be isolated and induced to exit the growth cycle and enter G₀ phase. The quiescent cell can then be fused, e.g., through the use of electrical pulses, to an enucleated oocyte from an animal of the same species from which the quiescent cell is isolated. The reconstructed oocyte is then cultured such that it develops to morula or blastocyte and then transferred to pseudopregnant female foster animal. The offspring borne of

this female foster animal will be a clone of the animal from which the cell, *e.g.*, the somatic cell, is isolated.

Pharmaceutical Compositions

The ORFX nucleic acid molecules, ORFX proteins, and anti-ORFX antibodies (also referred to herein as "active compounds") of the invention, and derivatives, fragments, analogs and homologs thereof, can be incorporated into pharmaceutical compositions suitable for administration. Such compositions typically comprise the nucleic acid molecule, protein, or antibody and a pharmaceutically acceptable carrier. As used herein, "pharmaceutically acceptable carrier" is intended to include any and all solvents, dispersion media, coatings, antibacterial and antifungal agents, isotonic and absorption delaying agents, and the like, compatible with pharmaceutical administration. Suitable carriers are described in the most recent edition of Remington's Pharmaceutical Sciences, a standard reference text in the field, which is incorporated herein by reference. Preferred examples of such carriers or diluents include, but are not limited to, water, saline, finger's solutions, dextrose solution, and 5% human serum albumin. Liposomes and non-aqueous vehicles such as fixed oils may also be used. The use of such media and agents for pharmaceutically active substances is well known in the art. Except insofar as any conventional media or agent is incompatible with the active compound, use thereof in the compositions is contemplated. Supplementary active compounds can also be incorporated into the compositions.

A pharmaceutical composition of the invention is formulated to be compatible with its intended route of administration. Examples of routes of administration include parenteral, *e.g.*, intravenous, intradermal, subcutaneous, oral (*e.g.*, inhalation), transdermal (topical), transmucosal, and rectal administration. Solutions or suspensions used for parenteral, intradermal, or subcutaneous application can include the following components: a sterile diluent such as water for injection, saline solution, fixed oils, polyethylene glycols, glycerine, propylene glycol or other synthetic solvents; antibacterial agents such as benzyl alcohol or methyl parabens; antioxidants such as ascorbic acid or sodium bisulfite; chelating agents such as ethylenediaminetetraacetic acid; buffers such as acetates, citrates or phosphates, and agents for the adjustment of tonicity such as sodium chloride or dextrose. The pH can be adjusted with acids or bases, such as hydrochloric acid or sodium hydroxide. The parenteral preparation can be enclosed in ampoules, disposable syringes or multiple dose vials made of glass or plastic.

Pharmaceutical compositions suitable for injectable use include sterile aqueous solutions (where water soluble) or dispersions and sterile powders for the extemporaneous preparation of sterile injectable solutions or dispersion. For intravenous administration, suitable carriers include physiological saline, bacteriostatic water, Cremophor EL™ (BASF, Parsippany, N.J.) or phosphate buffered saline (PBS). In all cases, the composition must be sterile and should be fluid to the extent that easy syringeability exists. It must be stable under the conditions of manufacture and storage and must be preserved against the contaminating action of microorganisms such as bacteria and fungi. The carrier can be a solvent or dispersion medium containing, for example, water, ethanol, polyol (for example, glycerol, propylene glycol, and liquid polyethylene glycol, and the like), and suitable mixtures thereof. The proper fluidity can be maintained, for example, by the use of a coating such as lecithin, by the maintenance of the required particle size in the case of dispersion and by the use of surfactants. Prevention of the action of microorganisms can be achieved by various antibacterial and antifungal agents, for example, parabens, chlorobutanol, phenol, ascorbic acid, thimerosal, and the like. In many cases, it will be preferable to include isotonic agents, for example, sugars, polyalcohols such as manitol, sorbitol, sodium chloride in the composition. Prolonged absorption of the injectable compositions can be brought about by including in the composition an agent which delays absorption, for example, aluminum monostearate and gelatin.

Sterile injectable solutions can be prepared by incorporating the active compound (*e.g.*, a ORFX protein or anti-ORFX antibody) in the required amount in an appropriate solvent with one or a combination of ingredients enumerated above, as required, followed by filtered sterilization. Generally, dispersions are prepared by incorporating the active compound into a sterile vehicle that contains a basic dispersion medium and the required other ingredients from those enumerated above. In the case of sterile powders for the preparation of sterile injectable solutions, methods of preparation are vacuum drying and freeze-drying that yields a powder of the active ingredient plus any additional desired ingredient from a previously sterile-filtered solution thereof.

Oral compositions generally include an inert diluent or an edible carrier. They can be enclosed in gelatin capsules or compressed into tablets. For the purpose of oral therapeutic administration, the active compound can be incorporated with excipients and used in the form of tablets, troches, or capsules. Oral compositions can also be prepared using a fluid carrier for use

as a mouthwash, wherein the compound in the fluid carrier is applied orally and swished and expectorated or swallowed. Pharmaceutically compatible binding agents, and/or adjuvant materials can be included as part of the composition. The tablets, pills, capsules, troches and the like can contain any of the following ingredients, or compounds of a similar nature: a binder
5 such as microcrystalline cellulose, gum tragacanth or gelatin; an excipient such as starch or lactose, a disintegrating agent such as alginic acid, Primogel, or corn starch; a lubricant such as magnesium stearate or Sterotes; a glidant such as colloidal silicon dioxide; a sweetening agent such as sucrose or saccharin; or a flavoring agent such as peppermint, methyl salicylate, or orange flavoring.

10 For administration by inhalation, the compounds are delivered in the form of an aerosol spray from pressured container or dispenser which contains a suitable propellant, *e.g.*, a gas such as carbon dioxide, or a nebulizer.

Systemic administration can also be by transmucosal or transdermal means. For transmucosal or transdermal administration, penetrants appropriate to the barrier to be permeated
15 are used in the formulation. Such penetrants are generally known in the art, and include, for example, for transmucosal administration, detergents, bile salts, and fusidic acid derivatives. Transmucosal administration can be accomplished through the use of nasal sprays or suppositories. For transdermal administration, the active compounds are formulated into ointments, salves, gels, or creams as generally known in the art.

20 The compounds can also be prepared in the form of suppositories (*e.g.*, with conventional suppository bases such as cocoa butter and other glycerides) or retention enemas for rectal delivery.

In one embodiment, the active compounds are prepared with carriers that will protect the compound against rapid elimination from the body, such as a controlled release formulation, including implants and microencapsulated delivery systems. Biodegradable, biocompatible
25 polymers can be used, such as ethylene vinyl acetate, polyanhydrides, polyglycolic acid, collagen, polyorthoesters, and polylactic acid. Methods for preparation of such formulations will be apparent to those skilled in the art. The materials can also be obtained commercially from Alza Corporation and Nova Pharmaceuticals, Inc. Liposomal suspensions (including liposomes targeted to infected cells with monoclonal antibodies to viral antigens) can also be used as
30 pharmaceutically acceptable carriers. These can be prepared according to methods known to those skilled in the art, for example, as described in U.S. Pat. No. 4,522,811.

It is especially advantageous to formulate oral or parenteral compositions in dosage unit form for ease of administration and uniformity of dosage. Dosage unit form as used herein refers to physically discrete units suited as unitary dosages for the subject to be treated; each unit containing a predetermined quantity of active compound calculated to produce the desired
5 therapeutic effect in association with the required pharmaceutical carrier. The specification for the dosage unit forms of the invention are dictated by and directly dependent on the unique characteristics of the active compound and the particular therapeutic effect to be achieved.

The nucleic acid molecules of the invention can be inserted into vectors and used as gene therapy vectors. Gene therapy vectors can be delivered to a subject by any of a number of
10 routes, *e.g.*, as described in U.S. Patent Nos. 5,703,055. Delivery can thus also include, *e.g.*, intravenous injection, local administration (see U.S. Pat. No. 5,328,470) or stereotactic injection (see *e.g.*, Chen *et al.* (1994) *PNAS* 91:3054-3057). The pharmaceutical preparation of the gene therapy vector can include the gene therapy vector in an acceptable diluent, or can comprise a slow release matrix in which the gene delivery vehicle is imbedded. Alternatively, where the
15 complete gene delivery vector can be produced intact from recombinant cells, *e.g.*, retroviral vectors, the pharmaceutical preparation can include one or more cells that produce the gene delivery system.

The pharmaceutical compositions can be included in a container, pack, or dispenser together with instructions for administration.

20 Additional Uses and Methods of the Invention

The nucleic acid molecules, proteins, protein homologues, and antibodies described herein can be used in one or more of the following methods: (a) screening assays; (b) detection assays (*e.g.*, chromosomal mapping, cell and tissue typing, forensic biology), (c) predictive medicine (*e.g.*, diagnostic assays, prognostic assays, monitoring clinical trials, and
25 pharmacogenomics); and (d) methods of treatment (*e.g.*, therapeutic and prophylactic).

The isolated nucleic acid molecules of the invention can be used to express ORFX protein (*e.g.*, via a recombinant expression vector in a host cell in gene therapy applications), to detect ORFX mRNA (*e.g.*, in a biological sample) or a genetic lesion in a ORFX gene, and to modulate ORFX activity, as described further below. In addition, the ORFX proteins can be
30 used to screen drugs or compounds that modulate the ORFX activity or expression as well as to treat disorders characterized by insufficient or excessive production of ORFX protein, for

example proliferative or differentiative disorders, or production of ORFX protein forms that have decreased or aberrant activity compared to ORFX wild type protein. In addition, the anti-ORFX antibodies of the invention can be used to detect and isolate ORFX proteins and modulate ORFX activity.

5 This invention further pertains to novel agents identified by the above described screening assays and uses thereof for treatments as described herein.

Screening Assays

 The invention provides a method (also referred to herein as a "screening assay") for identifying modulators, *i.e.*, candidate or test compounds or agents (*e.g.*, peptides,
10 peptidomimetics, small molecules or other drugs) that bind to ORFX proteins or have a stimulatory or inhibitory effect on, for example, ORFX expression or ORFX activity.

 In one embodiment, the invention provides assays for screening candidate or test compounds which bind to or modulate the activity of a ORFX protein or polypeptide or biologically active portion thereof. The test compounds of the present invention can be obtained
15 using any of the numerous approaches in combinatorial library methods known in the art, including: biological libraries; spatially addressable parallel solid phase or solution phase libraries; synthetic library methods requiring deconvolution; the "one-bead one-compound" library method; and synthetic library methods using affinity chromatography selection. The biological library approach is limited to peptide libraries, while the other four approaches are
20 applicable to peptide, non-peptide oligomer or small molecule libraries of compounds (Lam (1997) *Anticancer Drug Des* 12:145).

 Examples of methods for the synthesis of molecular libraries can be found in the art, for example in: DeWitt *et al.* (1993) *Proc Natl Acad Sci U.S.A.* 90:6909; Erb *et al.* (1994) *Proc Natl Acad Sci U.S.A.* 91:11422; Zuckermann *et al.* (1994) *J Med Chem* 37:2678; Cho *et al.* (1993)
25 *Science* 261:1303; Carrell *et al.* (1994) *Angew Chem Int Ed Engl* 33:2059; Carell *et al.* (1994) *Angew Chem Int Ed Engl* 33:2061; and Gallop *et al.* (1994) *J Med Chem* 37:1233.

 Libraries of compounds may be presented in solution (*e.g.*, Houghten (1992) *Biotechniques* 13:412-421), or on beads (Lam (1991) *Nature* 354:82-84), on chips (Fodor (1993) *Nature* 364:555-556), bacteria (Ladner U.S. Pat. No. 5,223,409), spores (Ladner USP '409),
30 plasmids (Cull *et al.* (1992) *Proc Natl Acad Sci USA* 89:1865-1869) or on phage (Scott and

Smith (1990) *Science* 249:386-390; Devlin (1990) *Science* 249:404-406; Cwirla *et al.* (1990) *Proc Natl Acad Sci U.S.A.* 87:6378-6382; Felici (1991) *J Mol Biol* 222:301-310; Ladner above.).

In one embodiment, an assay is a cell-based assay in which a cell which expresses a membrane-bound form of ORFX protein, or a biologically active portion thereof, on the cell surface is contacted with a test compound and the ability of the test compound to bind to a ORFX protein determined. The cell, for example, can be of mammalian origin or a yeast cell. Determining the ability of the test compound to bind to the ORFX protein can be accomplished, for example, by coupling the test compound with a radioisotope or enzymatic label such that binding of the test compound to the ORFX protein or biologically active portion thereof can be determined by detecting the labeled compound in a complex. For example, test compounds can be labeled with ^{125}I , ^{35}S , ^{14}C , or ^3H , either directly or indirectly, and the radioisotope detected by direct counting of radioemission or by scintillation counting. Alternatively, test compounds can be enzymatically labeled with, for example, horseradish peroxidase, alkaline phosphatase, or luciferase, and the enzymatic label detected by determination of conversion of an appropriate substrate to product. In one embodiment, the assay comprises contacting a cell which expresses a membrane-bound form of ORFX protein, or a biologically active portion thereof, on the cell surface with a known compound which binds ORFX to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to interact with a ORFX protein, wherein determining the ability of the test compound to interact with a ORFX protein comprises determining the ability of the test compound to preferentially bind to ORFX or a biologically active portion thereof as compared to the known compound.

In another embodiment, an assay is a cell-based assay comprising contacting a cell expressing a membrane-bound form of ORFX protein, or a biologically active portion thereof, on the cell surface with a test compound and determining the ability of the test compound to modulate (*e.g.*, stimulate or inhibit) the activity of the ORFX protein or biologically active portion thereof. Determining the ability of the test compound to modulate the activity of ORFX or a biologically active portion thereof can be accomplished, for example, by determining the ability of the ORFX protein to bind to or interact with a ORFX target molecule. As used herein, a "target molecule" is a molecule with which a ORFX protein binds or interacts in nature, for example, a molecule on the surface of a cell which expresses a ORFX interacting protein, a molecule on the surface of a second cell, a molecule in the extracellular milieu, a molecule

associated with the internal surface of a cell membrane or a cytoplasmic molecule. A ORFX target molecule can be a non-ORFX molecule or a ORFX protein or polypeptide of the present invention. In one embodiment, a ORFX target molecule is a component of a signal transduction pathway that facilitates transduction of an extracellular signal (*e.g.*, a signal generated by binding of a compound to a membrane-bound ORFX molecule) through the cell membrane and into the cell. The target, for example, can be a second intercellular protein that has catalytic activity or a protein that facilitates the association of downstream signaling molecules with ORFX.

Determining the ability of the ORFX protein to bind to or interact with a ORFX target molecule can be accomplished by one of the methods described above for determining direct binding. In one embodiment, determining the ability of the ORFX protein to bind to or interact with a ORFX target molecule can be accomplished by determining the activity of the target molecule. For example, the activity of the target molecule can be determined by detecting induction of a cellular second messenger of the target (*i.e.* intracellular Ca^{2+} , diacylglycerol, IP_3 , etc.), detecting catalytic/enzymatic activity of the target an appropriate substrate, detecting the induction of a reporter gene (comprising a ORFX-responsive regulatory element operatively linked to a nucleic acid encoding a detectable marker, *e.g.*, luciferase), or detecting a cellular response, for example, cell survival, cellular differentiation, or cell proliferation.

In yet another embodiment, an assay of the present invention is a cell-free assay comprising contacting a ORFX protein or biologically active portion thereof with a test compound and determining the ability of the test compound to bind to the ORFX protein or biologically active portion thereof. Binding of the test compound to the ORFX protein can be determined either directly or indirectly as described above. In one embodiment, the assay comprises contacting the ORFX protein or biologically active portion thereof with a known compound which binds ORFX to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to interact with a ORFX protein, wherein determining the ability of the test compound to interact with a ORFX protein comprises determining the ability of the test compound to preferentially bind to ORFX or biologically active portion thereof as compared to the known compound.

In another embodiment, an assay is a cell-free assay comprising contacting ORFX protein or biologically active portion thereof with a test compound and determining the ability of the test compound to modulate (*e.g.*, stimulate or inhibit) the activity of the ORFX protein or

biologically active portion thereof. Determining the ability of the test compound to modulate the activity of ORFX can be accomplished, for example, by determining the ability of the ORFX protein to bind to a ORFX target molecule by one of the methods described above for determining direct binding. In an alternative embodiment, determining the ability of the test compound to modulate the activity of ORFX can be accomplished by determining the ability of the ORFX protein further modulate a ORFX target molecule. For example, the catalytic/enzymatic activity of the target molecule on an appropriate substrate can be determined as previously described.

In yet another embodiment, the cell-free assay comprises contacting the ORFX protein or biologically active portion thereof with a known compound which binds ORFX to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to interact with a ORFX protein, wherein determining the ability of the test compound to interact with a ORFX protein comprises determining the ability of the ORFX protein to preferentially bind to or modulate the activity of a ORFX target molecule.

The cell-free assays of the present invention are amenable to use of both the soluble form or the membrane-bound form of ORFX. In the case of cell-free assays comprising the membrane-bound form of ORFX, it may be desirable to utilize a solubilizing agent such that the membrane-bound form of ORFX is maintained in solution. Examples of such solubilizing agents include non-ionic detergents such as n-octylglucoside, n-dodecylglucoside, n-dodecylmaltoside, octanoyl-N-methylglucamide, decanoyl-N-methylglucamide, Triton[®] X-100, Triton[®] X-114, Thesit[®], Isotridecypoly(ethylene glycol ether)_n, N-dodecyl--N,N-dimethyl-3-ammonio-1-propane sulfonate, 3-(3-cholamidopropyl)dimethylamminiol-1-propane sulfonate (CHAPS), or 3-(3-cholamidopropyl)dimethylamminiol-2-hydroxy-1-propane sulfonate (CHAPSO).

In more than one embodiment of the above assay methods of the present invention, it may be desirable to immobilize either ORFX or its target molecule to facilitate separation of complexed from uncomplexed forms of one or both of the proteins, as well as to accommodate automation of the assay. Binding of a test compound to ORFX, or interaction of ORFX with a target molecule in the presence and absence of a candidate compound, can be accomplished in any vessel suitable for containing the reactants. Examples of such vessels include microtiter plates, test tubes, and micro-centrifuge tubes. In one embodiment, a fusion protein can be provided that adds a domain that allows one or both of the proteins to be bound to a matrix. For

example, GST-ORFX fusion proteins or GST-target fusion proteins can be adsorbed onto glutathione sepharose beads (Sigma Chemical, St. Louis, MO) or glutathione derivatized microtiter plates, that are then combined with the test compound or the test compound and either the non-adsorbed target protein or ORFX protein, and the mixture is incubated under conditions conducive to complex formation (*e.g.*, at physiological conditions for salt and pH). Following incubation, the beads or microtiter plate wells are washed to remove any unbound components, the matrix immobilized in the case of beads, complex determined either directly or indirectly, for example, as described above. Alternatively, the complexes can be dissociated from the matrix, and the level of ORFX binding or activity determined using standard techniques.

Other techniques for immobilizing proteins on matrices can also be used in the screening assays of the invention. For example, either ORFX or its target molecule can be immobilized utilizing conjugation of biotin and streptavidin. Biotinylated ORFX or target molecules can be prepared from biotin-NHS (N-hydroxy-succinimide) using techniques well known in the art (*e.g.*, biotinylation kit, Pierce Chemicals, Rockford, Ill.), and immobilized in the wells of streptavidin-coated 96 well plates (Pierce Chemical). Alternatively, antibodies reactive with ORFX or target molecules, but which do not interfere with binding of the ORFX protein to its target molecule, can be derivatized to the wells of the plate, and unbound target or ORFX trapped in the wells by antibody conjugation. Methods for detecting such complexes, in addition to those described above for the GST-immobilized complexes, include immunodetection of complexes using antibodies reactive with the ORFX or target molecule, as well as enzyme-linked assays that rely on detecting an enzymatic activity associated with the ORFX or target molecule.

In another embodiment, modulators of ORFX expression are identified in a method wherein a cell is contacted with a candidate compound and the expression of ORFX mRNA or protein in the cell is determined. The level of expression of ORFX mRNA or protein in the presence of the candidate compound is compared to the level of expression of ORFX mRNA or protein in the absence of the candidate compound. The candidate compound can then be identified as a modulator of ORFX expression based on this comparison. For example, when expression of ORFX mRNA or protein is greater (statistically significantly greater) in the presence of the candidate compound than in its absence, the candidate compound is identified as a stimulator of ORFX mRNA or protein expression. Alternatively, when expression of ORFX mRNA or protein is less (statistically significantly less) in the presence of the candidate

compound than in its absence, the candidate compound is identified as an inhibitor of ORFX mRNA or protein expression. The level of ORFX mRNA or protein expression in the cells can be determined by methods described herein for detecting ORFX mRNA or protein.

In yet another aspect of the invention, the ORFX proteins can be used as "bait proteins" in a two-hybrid assay or three hybrid assay (see, *e.g.*, U.S. Pat. No. 5,283,317; Zervos *et al.* (1993) Cell 72:223-232; Madura *et al.* (1993) J Biol Chem 268:12046-12054; Bartel *et al.* (1993) Biotechniques 14:920-924; Iwabuchi *et al.* (1993) Oncogene 8:1693-1696; and Brent WO94/10300), to identify other proteins that bind to or interact with ORFX ("ORFX-binding proteins" or "ORFX-bp") and modulate ORFX activity. Such ORFX-binding proteins are also likely to be involved in the propagation of signals by the ORFX proteins as, for example, upstream or downstream elements of the ORFX pathway.

The two-hybrid system is based on the modular nature of most transcription factors, which consist of separable DNA-binding and activation domains. Briefly, the assay utilizes two different DNA constructs. In one construct, the gene that codes for ORFX is fused to a gene encoding the DNA binding domain of a known transcription factor (*e.g.*, GAL-4). In the other construct, a DNA sequence, from a library of DNA sequences, that encodes an unidentified protein ("prey" or "sample") is fused to a gene that codes for the activation domain of the known transcription factor. If the "bait" and the "prey" proteins are able to interact, *in vivo*, forming a ORFX-dependent complex, the DNA-binding and activation domains of the transcription factor are brought into close proximity. This proximity allows transcription of a reporter gene (*e.g.*, LacZ) that is operably linked to a transcriptional regulatory site responsive to the transcription factor. Expression of the reporter gene can be detected and cell colonies containing the functional transcription factor can be isolated and used to obtain the cloned gene that encodes the protein which interacts with ORFX.

This invention further pertains to novel agents identified by the above-described screening assays and uses thereof for treatments as described herein.

Detection Assays

Portions or fragments of the cDNA sequences identified herein (and the corresponding complete gene sequences) can be used in numerous ways as polynucleotide reagents. For example, these sequences can be used to: (i) map their respective genes on a chromosome; and, thus, locate gene regions associated with genetic disease; (ii) identify an individual from a

minute biological sample (tissue typing); and (iii) aid in forensic identification of a biological sample.

The ORFX sequences of the present invention can also be used to identify individuals from minute biological samples. In this technique, an individual's genomic DNA is digested with one or more restriction enzymes, and probed on a Southern blot to yield unique bands for identification. The sequences of the present invention are useful as additional DNA markers for RFLP ("restriction fragment length polymorphisms," described in U.S. Pat. No. 5,272,057).

Furthermore, the sequences of the present invention can be used to provide an alternative technique that determines the actual base-by-base DNA sequence of selected portions of an individual's genome. Thus, the ORFX sequences described herein can be used to prepare two PCR primers from the 5' and 3' ends of the sequences. These primers can then be used to amplify an individual's DNA and subsequently sequence it.

Panels of corresponding DNA sequences from individuals, prepared in this manner, can provide unique individual identifications, as each individual will have a unique set of such DNA sequences due to allelic differences. The sequences of the present invention can be used to obtain such identification sequences from individuals and from tissue. The ORFX sequences of the invention uniquely represent portions of the human genome. Allelic variation occurs to some degree in the coding regions of these sequences, and to a greater degree in the noncoding regions. It is estimated that allelic variation between individual humans occurs with a frequency of about once per each 500 bases. Much of the allelic variation is due to single nucleotide polymorphisms (SNPs), which include restriction fragment length polymorphisms (RFLPs).

Each of the sequences described herein can, to some degree, be used as a standard against which DNA from an individual can be compared for identification purposes. Because greater numbers of polymorphisms occur in the noncoding regions, fewer sequences are necessary to differentiate individuals. The noncoding sequences of SEQ ID NO:2 n -1 (wherein $n = 1$ to 3161), as described above, can comfortably provide positive individual identification with a panel of perhaps 10 to 1,000 primers that each yield a noncoding amplified sequence of 100 bases. If predicted coding sequences are used, a more appropriate number of primers for positive individual identification would be 500-2,000.

Predictive Medicine

The present invention also pertains to the field of predictive medicine in which diagnostic assays, prognostic assays, pharmacogenomics, and monitoring clinical trials are used for prognostic (predictive) purposes to thereby treat an individual prophylactically. Accordingly, one aspect of the present invention relates to diagnostic assays for determining ORFX protein and/or nucleic acid expression as well as ORFX activity, in the context of a biological sample (*e.g.*, blood, serum, cells, tissue) to thereby determine whether an individual is afflicted with a disease or disorder, or is at risk of developing a disorder, associated with aberrant ORFX expression or activity. The invention also provides for prognostic (or predictive) assays for determining whether an individual is at risk of developing a disorder associated with ORFX protein, nucleic acid expression or activity. For example, mutations in a ORFX gene can be assayed in a biological sample. Such assays can be used for prognostic or predictive purpose to thereby prophylactically treat an individual prior to the onset of a disorder characterized by or associated with ORFX protein, nucleic acid expression or activity.

Another aspect of the invention provides methods for determining ORFX protein, nucleic acid expression or ORFX activity in an individual to thereby select appropriate therapeutic or prophylactic agents for that individual (referred to herein as "pharmacogenomics"). Pharmacogenomics allows for the selection of agents (*e.g.*, drugs) for therapeutic or prophylactic treatment of an individual based on the genotype of the individual (*e.g.*, the genotype of the individual examined to determine the ability of the individual to respond to a particular agent.)

Yet another aspect of the invention pertains to monitoring the influence of agents (*e.g.*, drugs, compounds) on the expression or activity of ORFX in clinical trials.

Use of Partial ORFX Sequences in Forensic Biology

DNA-based identification techniques can also be used in forensic biology. Forensic biology is a scientific field employing genetic typing of biological evidence found at a crime scene as a means for positively identifying, for example, a perpetrator of a crime. To make such an identification, PCR technology can be used to amplify DNA sequences taken from very small biological samples such as tissues, *e.g.*, hair or skin, or body fluids, *e.g.*, blood, saliva, or semen

found at a crime scene. The amplified sequence can then be compared to a standard, thereby allowing identification of the origin of the biological sample.

The sequences of the present invention can be used to provide polynucleotide reagents, *e.g.*, PCR primers, targeted to specific loci in the human genome, that can enhance the reliability of DNA-based forensic identifications by, for example, providing another "identification marker" (*i.e.* another DNA sequence that is unique to a particular individual). As mentioned above, actual base sequence information can be used for identification as an accurate alternative to patterns formed by restriction enzyme generated fragments. Sequences targeted to noncoding regions of SEQ ID NOs: __ are particularly appropriate for this use as greater numbers of polymorphisms occur in the noncoding regions, making it easier to differentiate individuals using this technique. Examples of polynucleotide reagents include the ORFX sequences or portions thereof, *e.g.*, fragments derived from the noncoding regions of one or more of SEQ ID NO:2*n*-1 (where *n* = 1 to 3161), having a length of at least 20 bases, preferably at least 30 bases.

The ORFX sequences described herein can further be used to provide polynucleotide reagents, *e.g.*, labeled or label-able probes that can be used, for example, in an *in situ* hybridization technique, to identify a specific tissue, *e.g.*, brain tissue, etc. This can be very useful in cases where a forensic pathologist is presented with a tissue of unknown origin. Panels of such ORFX probes can be used to identify tissue by species and/or by organ type.

In a similar fashion, these reagents, *e.g.*, ORFX primers or probes can be used to screen tissue culture for contamination (*i.e.* screen for the presence of a mixture of different types of cells in a culture).

Predictive Medicine

The present invention also pertains to the field of predictive medicine in which diagnostic assays, prognostic assays, pharmacogenomics, and monitoring clinical trials are used for prognostic (predictive) purposes to thereby treat an individual prophylactically. Accordingly, one aspect of the present invention relates to diagnostic assays for determining ORFX protein and/or nucleic acid expression as well as ORFX activity, in the context of a biological sample (*e.g.*, blood, serum, cells, tissue) to thereby determine whether an individual is afflicted with a disease or disorder, or is at risk of developing a disorder, associated with aberrant ORFX expression or activity. The invention also provides for prognostic (or predictive) assays for determining whether an individual is at risk of developing a disorder associated with ORFX protein, nucleic

acid expression or activity. For example, mutations in a ORFX gene can be assayed in a biological sample. Such assays can be used for prognostic or predictive purpose to thereby prophylactically treat an individual prior to the onset of a disorder characterized by or associated with ORFX protein, nucleic acid expression or activity.

5 Another aspect of the invention provides methods for determining ORFX protein, nucleic acid expression or ORFX activity in an individual to thereby select appropriate therapeutic or prophylactic agents for that individual (referred to herein as "pharmacogenomics"). Pharmacogenomics allows for the selection of agents (*e.g.*, drugs) for therapeutic or prophylactic treatment of an individual based on the genotype of the individual (*e.g.*, the genotype of the individual examined to determine the ability of the individual to respond to a particular agent.)

10 Yet another aspect of the invention pertains to monitoring the influence of agents (*e.g.*, drugs, compounds) on the expression or activity of ORFX in clinical trials.

These and other agents are described in further detail in the following sections.

Diagnostic Assays

15 Other conditions in which proliferation of cells plays a role include tumors, restenosis, psoriasis, Dupuytren's contracture, diabetic complications, Kaposi's sarcoma and rheumatoid arthritis.

20 An ORFX polypeptide may be used to identify an interacting polypeptide a sample or tissue. The method comprises contacting the sample or tissue with ORFX, allowing formation of a complex between the ORFX polypeptide and the interacting polypeptide, and detecting the complex, if present.

25 The proteins of the invention may be used to stimulate production of antibodies specifically binding the proteins. Such antibodies may be used in immunodiagnostic procedures to detect the occurrence of the protein in a sample. The proteins of the invention may be used to stimulate cell growth and cell proliferation in conditions in which such growth would be favorable. An example would be to counteract toxic side effects of chemotherapeutic agents on, for example, hematopoiesis and platelet formation, linings of the gastrointestinal tract, and hair follicles. They may also be used to stimulate new cell growth in neurological disorders including, for example, Alzheimer's disease. Alternatively, antagonistic treatments may be

30 administered in which an antibody specifically binding the ORFX-like proteins of the invention

would abrogate the specific growth-inducing effects of the proteins. Such antibodies may be useful, for example, in the treatment of proliferative disorders including various tumors and benign hyperplasias.

Polynucleotides or oligonucleotides corresponding to any one portion of the ORFX nucleic acids of SEQ ID NO:2n-1 (wherein $n = 1$ to 3161) may be used to detect DNA containing a corresponding ORF gene, or detect the expression of a corresponding ORFX gene, or ORFX-like gene. For example, an ORFX nucleic acid expressed in a particular cell or tissue, as noted in Table 2, can be used to identify the presence of that particular cell type.

An exemplary method for detecting the presence or absence of ORFX in a biological sample involves obtaining a biological sample from a test subject and contacting the biological sample with a compound or an agent capable of detecting ORFX protein or nucleic acid (*e.g.*, mRNA, genomic DNA) that encodes ORFX protein such that the presence of ORFX is detected in the biological sample. An agent for detecting ORFX mRNA or genomic DNA is a labeled nucleic acid probe capable of hybridizing to ORFX mRNA or genomic DNA. The nucleic acid probe can be, for example, a full-length ORFX nucleic acid, such as the nucleic acid of SEQ ID NO:2n-1 (wherein $n = 1$ to 3161), or a portion thereof, such as an oligonucleotide of at least 15, 30, 50, 100, 250 or 500 nucleotides in length and sufficient to specifically hybridize under stringent conditions to ORFX mRNA or genomic DNA, as described above. Other suitable probes for use in the diagnostic assays of the invention are described herein.

An agent for detecting ORFX protein is an antibody capable of binding to ORFX protein, preferably an antibody with a detectable label. Antibodies can be polyclonal, or more preferably, monoclonal. An intact antibody, or a fragment thereof (*e.g.*, Fab or F(ab')₂) can be used. The term "labeled", with regard to the probe or antibody, is intended to encompass direct labeling of the probe or antibody by coupling (*i.e.*, physically linking) a detectable substance to the probe or antibody, as well as indirect labeling of the probe or antibody by reactivity with another reagent that is directly labeled. Examples of indirect labeling include detection of a primary antibody using a fluorescently labeled secondary antibody and end-labeling of a DNA probe with biotin such that it can be detected with fluorescently labeled streptavidin. The term "biological sample" is intended to include tissues, cells and biological fluids isolated from a subject, as well as tissues, cells and fluids present within a subject. That is, the detection method of the invention can be used to detect ORFX mRNA, protein, or genomic DNA in a biological sample *in vitro* as

well as *in vivo*. For example, *in vitro* techniques for detection of ORFX mRNA include Northern hybridizations and *in situ* hybridizations. *In vitro* techniques for detection of ORFX protein include enzyme linked immunosorbent assays (ELISAs), Western blots, immunoprecipitations and immunofluorescence. *In vitro* techniques for detection of ORFX genomic DNA include Southern hybridizations. Furthermore, *in vivo* techniques for detection of ORFX protein include introducing into a subject a labeled anti-ORFX antibody. For example, the antibody can be labeled with a radioactive marker whose presence and location in a subject can be detected by standard imaging techniques.

In one embodiment, the biological sample contains protein molecules from the test subject. Alternatively, the biological sample can contain mRNA molecules from the test subject or genomic DNA molecules from the test subject. A preferred biological sample is a peripheral blood leukocyte sample isolated by conventional means from a subject.

In another embodiment, the methods further involve obtaining a control biological sample from a control subject, contacting the control sample with a compound or agent capable of detecting ORFX protein, mRNA, or genomic DNA, such that the presence of ORFX protein, mRNA or genomic DNA is detected in the biological sample, and comparing the presence of ORFX protein, mRNA or genomic DNA in the control sample with the presence of ORFX protein, mRNA or genomic DNA in the test sample.

The invention also encompasses kits for detecting the presence of ORFX in a biological sample. For example, the kit can comprise: a labeled compound or agent capable of detecting ORFX protein or mRNA in a biological sample; means for determining the amount of ORFX in the sample; and means for comparing the amount of ORFX in the sample with a standard. The compound or agent can be packaged in a suitable container. The kit can further comprise instructions for using the kit to detect ORFX protein or nucleic acid.

Prognostic Assays

The diagnostic methods described herein can furthermore be utilized to identify subjects having or at risk of developing a disease or disorder associated with aberrant ORFX expression or activity. For example, the assays described herein, such as the preceding diagnostic assays or the following assays, can be utilized to identify a subject having or at risk of developing a disorder associated with ORFX protein, nucleic acid expression or activity in, *e.g.*, proliferative or differentiative disorders such as hyperplasias, tumors, restenosis, psoriasis, Dupuytren's

contracture, diabetic complications, or rheumatoid arthritis, etc.; and glia-associated disorders such as cerebral lesions, diabetic neuropathies, cerebral edema, senile dementia, Alzheimer's disease, etc. Alternatively, the prognostic assays can be utilized to identify a subject having or at risk for developing a disease or disorder. Thus, the present invention provides a method for identifying a disease or disorder associated with aberrant ORFX expression or activity in which a test sample is obtained from a subject and ORFX protein or nucleic acid (*e.g.*, mRNA, genomic DNA) is detected, wherein the presence of ORFX protein or nucleic acid is diagnostic for a subject having or at risk of developing a disease or disorder associated with aberrant ORFX expression or activity. As used herein, a "test sample" refers to a biological sample obtained from a subject of interest. For example, a test sample can be a biological fluid (*e.g.*, serum), cell sample, or tissue.

Furthermore, the prognostic assays described herein can be used to determine whether a subject can be administered an agent (*e.g.*, an agonist, antagonist, peptidomimetic, protein, peptide, nucleic acid, small molecule, or other drug candidate) to treat a disease or disorder associated with aberrant ORFX expression or activity. For example, such methods can be used to determine whether a subject can be effectively treated with an agent for a disorder, such as a proliferative disorder, differentiative disorder, glia-associated disorders, etc. Thus, the present invention provides methods for determining whether a subject can be effectively treated with an agent for a disorder associated with aberrant ORFX expression or activity in which a test sample is obtained and ORFX protein or nucleic acid is detected (*e.g.*, wherein the presence of ORFX protein or nucleic acid is diagnostic for a subject that can be administered the agent to treat a disorder associated with aberrant ORFX expression or activity.)

The methods of the invention can also be used to detect genetic lesions in a ORFX gene, thereby determining if a subject with the lesioned gene is at risk for, or suffers from, a proliferative disorder, differentiative disorder, glia-associated disorder, etc. In various embodiments, the methods include detecting, in a sample of cells from the subject, the presence or absence of a genetic lesion characterized by at least one of an alteration affecting the integrity of a gene encoding a ORFX-protein, or the mis-expression of the ORFX gene. For example, such genetic lesions can be detected by ascertaining the existence of at least one of (1) a deletion of one or more nucleotides from a ORFX gene; (2) an addition of one or more nucleotides to a ORFX gene; (3) a substitution of one or more nucleotides of a ORFX gene, (4) a chromosomal

rearrangement of a ORFX gene; (5) an alteration in the level of a messenger RNA transcript of a ORFX gene, (6) aberrant modification of a ORFX gene, such as of the methylation pattern of the genomic DNA, (7) the presence of a non-wild type splicing pattern of a messenger RNA transcript of a ORFX gene, (8) a non-wild type level of a ORFX-protein, (9) allelic loss of a ORFX gene, and (10) inappropriate post-translational modification of a ORFX-protein. As described herein, there are a large number of assay techniques known in the art which can be used for detecting lesions in a ORFX gene. A preferred biological sample is a peripheral blood leukocyte sample isolated by conventional means from a subject. However, any biological sample containing nucleated cells may be used, including, for example, buccal mucosal cells.

In certain embodiments, detection of the lesion involves the use of a probe/primer in a polymerase chain reaction (PCR) (see, e.g., U.S. Pat. Nos. 4,683,195 and 4,683,202), such as anchor PCR or RACE PCR, or, alternatively, in a ligation chain reaction (LCR) (see, e.g., Landegran *et al.* (1988) *Science* 241:1077-1080; and Nakazawa *et al.* (1994) *PNAS* 91:360-364), the latter of which can be particularly useful for detecting point mutations in the ORFX-gene (see Abravaya *et al.* (1995) *Nucl Acids Res* 23:675-682). This method can include the steps of collecting a sample of cells from a patient, isolating nucleic acid (e.g., genomic, mRNA or both) from the cells of the sample, contacting the nucleic acid sample with one or more primers that specifically hybridize to a ORFX gene under conditions such that hybridization and amplification of the ORFX gene (if present) occurs, and detecting the presence or absence of an amplification product, or detecting the size of the amplification product and comparing the length to a control sample. It is anticipated that PCR and/or LCR may be desirable to use as a preliminary amplification step in conjunction with any of the techniques used for detecting mutations described herein.

Alternative amplification methods include: self sustained sequence replication (Guatelli *et al.*, 1990, *Proc Natl Acad Sci USA* 87:1874-1878), transcriptional amplification system (Kwoh, *et al.*, 1989, *Proc Natl Acad Sci USA* 86:1173-1177), Q-Beta Replicase (Lizardi *et al.*, 1988, *BioTechnology* 6:1197), or any other nucleic acid amplification method, followed by the detection of the amplified molecules using techniques well known to those of skill in the art. These detection schemes are especially useful for the detection of nucleic acid molecules if such molecules are present in very low numbers.

In an alternative embodiment, mutations in a ORFX gene from a sample cell can be identified by alterations in restriction enzyme cleavage patterns. For example, sample and control DNA is isolated, amplified (optionally), digested with one or more restriction endonucleases, and fragment length sizes are determined by gel electrophoresis and compared. Differences in fragment length sizes between sample and control DNA indicates mutations in the sample DNA. Moreover, the use of sequence specific ribozymes (see, for example, U.S. Pat. No. 5,493,531) can be used to score for the presence of specific mutations by development or loss of a ribozyme cleavage site.

In other embodiments, genetic mutations in ORFX can be identified by hybridizing a sample and control nucleic acids, *e.g.*, DNA or RNA, to high density arrays containing hundreds or thousands of oligonucleotides probes (Cronin *et al.* (1996) *Human Mutation* 7: 244-255; Kozal *et al.* (1996) *Nature Medicine* 2: 753-759). For example, genetic mutations in ORFX can be identified in two dimensional arrays containing light-generated DNA probes as described in Cronin *et al.* above. Briefly, a first hybridization array of probes can be used to scan through long stretches of DNA in a sample and control to identify base changes between the sequences by making linear arrays of sequential overlapping probes. This step allows the identification of point mutations. This step is followed by a second hybridization array that allows the characterization of specific mutations by using smaller, specialized probe arrays complementary to all variants or mutations detected. Each mutation array is composed of parallel probe sets, one complementary to the wild-type gene and the other complementary to the mutant gene.

In yet another embodiment, any of a variety of sequencing reactions known in the art can be used to directly sequence the ORFX gene and detect mutations by comparing the sequence of the sample ORFX with the corresponding wild-type (control) sequence. Examples of sequencing reactions include those based on techniques developed by Maxim and Gilbert (1977) *PNAS* 74:560 or Sanger (1977) *PNAS* 74:5463. It is also contemplated that any of a variety of automated sequencing procedures can be utilized when performing the diagnostic assays (Naeve *et al.*, (1995) *Biotechniques* 19:448), including sequencing by mass spectrometry (see, *e.g.*, PCT International Publ. No. WO 94/16101; Cohen *et al.* (1996) *Adv Chromatogr* 36:127-162; and Griffin *et al.* (1993) *Appl Biochem Biotechnol* 38:147-159).

Other methods for detecting mutations in the ORFX gene include methods in which protection from cleavage agents is used to detect mismatched bases in RNA/RNA or RNA/DNA

heteroduplexes (Myers *et al.* (1985) *Science* 230:1242). In general, the art technique of "mismatch cleavage" starts by providing heteroduplexes of formed by hybridizing (labeled) RNA or DNA containing the wild-type ORFX sequence with potentially mutant RNA or DNA obtained from a tissue sample. The double-stranded duplexes are treated with an agent that cleaves single-stranded regions of the duplex such as which will exist due to basepair mismatches between the control and sample strands. For instance, RNA/DNA duplexes can be treated with RNase and DNA/DNA hybrids treated with S1 nuclease to enzymatically digesting the mismatched regions. In other embodiments, either DNA/DNA or RNA/DNA duplexes can be treated with hydroxylamine or osmium tetroxide and with piperidine in order to digest mismatched regions. After digestion of the mismatched regions, the resulting material is then separated by size on denaturing polyacrylamide gels to determine the site of mutation. See, for example, Cotton *et al* (1988) *Proc Natl Acad Sci USA* 85:4397; Saleeba *et al* (1992) *Methods Enzymol* 217:286-295. In an embodiment, the control DNA or RNA can be labeled for detection.

In still another embodiment, the mismatch cleavage reaction employs one or more proteins that recognize mismatched base pairs in double-stranded DNA (so called "DNA mismatch repair" enzymes) in defined systems for detecting and mapping point mutations in ORFX cDNAs obtained from samples of cells. For example, the mutY enzyme of *E. coli* cleaves A at G/A mismatches and the thymidine DNA glycosylase from HeLa cells cleaves T at G/T mismatches (Hsu *et al.* (1994) *Carcinogenesis* 15:1657-1662). According to an exemplary embodiment, a probe based on a ORFX sequence, *e.g.*, a wild-type ORFX sequence, is hybridized to a cDNA or other DNA product from a test cell(s). The duplex is treated with a DNA mismatch repair enzyme, and the cleavage products, if any, can be detected from electrophoresis protocols or the like. See, for example, U.S. Pat. No. 5,459,039.

In other embodiments, alterations in electrophoretic mobility will be used to identify mutations in ORFX genes. For example, single strand conformation polymorphism (SSCP) may be used to detect differences in electrophoretic mobility between mutant and wild type nucleic acids (Orita *et al.* (1989) *Proc Natl Acad Sci USA*: 86:2766, see also Cotton (1993) *Mutat Res* 285:125-144; Hayashi (1992) *Genet Anal Tech Appl* 9:73-79). Single-stranded DNA fragments of sample and control ORFX nucleic acids will be denatured and allowed to renature. The secondary structure of single-stranded nucleic acids varies according to sequence, the resulting

alteration in electrophoretic mobility enables the detection of even a single base change. The DNA fragments may be labeled or detected with labeled probes. The sensitivity of the assay may be enhanced by using RNA, rather than DNA, in which the secondary structure is more sensitive to a change in sequence. In one embodiment, the subject method utilizes heteroduplex analysis to separate double stranded heteroduplex molecules on the basis of changes in electrophoretic mobility. See, e.g., Keen *et al.* (1991) *Trends Genet* 7:5.

In yet another embodiment the movement of mutant or wild-type fragments in polyacrylamide gels containing a gradient of denaturant is assayed using denaturing gradient gel electrophoresis (DGGE). See, e.g., Myers *et al.* (1985) *Nature* 313:495. When DGGE is used as the method of analysis, DNA will be modified to insure that it does not completely denature, for example by adding a GC clamp of approximately 40 bp of high-melting GC-rich DNA by PCR. In a further embodiment, a temperature gradient is used in place of a denaturing gradient to identify differences in the mobility of control and sample DNA. See, e.g., Rosenbaum and Reissner (1987) *Biophys Chem* 265:12753.

Examples of other techniques for detecting point mutations include, but are not limited to, selective oligonucleotide hybridization, selective amplification, or selective primer extension. For example, oligonucleotide primers may be prepared in which the known mutation is placed centrally and then hybridized to target DNA under conditions that permit hybridization only if a perfect match is found. See, e.g., Saiki *et al.* (1986) *Nature* 324:163; Saiki *et al.* (1989) *Proc Natl Acad. Sci USA* 86:6230. Such allele specific oligonucleotides are hybridized to PCR amplified target DNA or a number of different mutations when the oligonucleotides are attached to the hybridizing membrane and hybridized with labeled target DNA.

Alternatively, allele specific amplification technology that depends on selective PCR amplification may be used in conjunction with the instant invention. Oligonucleotides used as primers for specific amplification may carry the mutation of interest in the center of the molecule (so that amplification depends on differential hybridization) (Gibbs *et al.* (1989) *Nucleic Acids Res* 17:2437-2448) or at the extreme 3' end of one primer where, under appropriate conditions, mismatch can prevent, or reduce polymerase extension (Prossner (1993) *Tibtech* 11:238). In addition it may be desirable to introduce a novel restriction site in the region of the mutation to create cleavage-based detection. See, e.g., Gasparini *et al.* (1992) *Mol Cell Probes* 6:1. It is anticipated that in certain embodiments amplification may also be performed using Taq ligase

for amplification. See, *e.g.*, Barany (1991) *Proc Natl Acad Sci USA* 88:189. In such cases, ligation will occur only if there is a perfect match at the 3' end of the 5' sequence, making it possible to detect the presence of a known mutation at a specific site by looking for the presence or absence of amplification.

5 The methods described herein may be performed, for example, by utilizing pre-packaged diagnostic kits comprising at least one probe nucleic acid or antibody reagent described herein, which may be conveniently used, *e.g.*, in clinical settings to diagnose patients exhibiting symptoms or family history of a disease or illness involving a ORFX gene.

10 Furthermore, any cell type or tissue, preferably peripheral blood leukocytes, in which ORFX is expressed may be utilized in the prognostic assays described herein. However, any biological sample containing nucleated cells may be used, including, for example, buccal mucosal cells.

Pharmacogenomics

15 Agents, or modulators that have a stimulatory or inhibitory effect on ORFX activity (*e.g.*, ORFX gene expression), as identified by a screening assay described herein can be administered to individuals to treat (prophylactically or therapeutically) disorders (*e.g.*, neurological, cancer-related or gestational disorders) associated with aberrant ORFX activity. In conjunction with such treatment, the pharmacogenomics (*i.e.*, the study of the relationship between an individual's genotype and that individual's response to a foreign compound or drug) of the individual may be
20 considered. Differences in metabolism of therapeutics can lead to severe toxicity or therapeutic failure by altering the relation between dose and blood concentration of the pharmacologically active drug. Thus, the pharmacogenomics of the individual permits the selection of effective agents (*e.g.*, drugs) for prophylactic or therapeutic treatments based on a consideration of the individual's genotype. Such pharmacogenomics can further be used to determine appropriate
25 dosages and therapeutic regimens. Accordingly, the activity of ORFX protein, expression of ORFX nucleic acid, or mutation content of ORFX genes in an individual can be determined to thereby select appropriate agent(s) for therapeutic or prophylactic treatment of the individual.

30 Pharmacogenomics deals with clinically significant hereditary variations in the response to drugs due to altered drug disposition and abnormal action in affected persons. See *e.g.*, Eichelbaum, 1996, *Clin Exp Pharmacol Physiol*, 23:983-985 and Linder, 1997, *Clin Chem*, 43:254-266. In general, two types of pharmacogenetic conditions can be differentiated. Genetic

conditions transmitted as a single factor altering the way drugs act on the body (altered drug action) or genetic conditions transmitted as single factors altering the way the body acts on drugs (altered drug metabolism). These pharmacogenetic conditions can occur either as rare defects or as polymorphisms. For example, glucose-6-phosphate dehydrogenase (G6PD) deficiency is a common inherited enzymopathy in which the main clinical complication is haemolysis after ingestion of oxidant drugs (anti-malarials, sulfonamides, analgesics, nitrofurans) and consumption of fava beans.

As an illustrative embodiment, the activity of drug metabolizing enzymes is a major determinant of both the intensity and duration of drug action. The discovery of genetic polymorphisms of drug metabolizing enzymes (*e.g.*, N-acetyltransferase 2 (NAT 2) and cytochrome P450 enzymes CYP2D6 and CYP2C19) has provided an explanation as to why some patients do not obtain the expected drug effects or show exaggerated drug response and serious toxicity after taking the standard and safe dose of a drug. These polymorphisms are expressed in two phenotypes in the population, the extensive metabolizer (EM) and poor metabolizer (PM). The prevalence of PM is different among different populations. For example, the gene coding for CYP2D6 is highly polymorphic and several mutations have been identified in PM, which all lead to the absence of functional CYP2D6. Poor metabolizers of CYP2D6 and CYP2C19 quite frequently experience exaggerated drug response and side effects when they receive standard doses. If a metabolite is the active therapeutic moiety, PM show no therapeutic response, as demonstrated for the analgesic effect of codeine mediated by its CYP2D6-formed metabolite morphine. The other extreme are the so called ultra-rapid metabolizers who do not respond to standard doses. Recently, the molecular basis of ultra-rapid metabolism has been identified to be due to CYP2D6 gene amplification.

Thus, the activity of ORFX protein, expression of ORFX nucleic acid, or mutation content of ORFX genes in an individual can be determined to thereby select appropriate agent(s) for therapeutic or prophylactic treatment of the individual. In addition, pharmacogenetic studies can be used to apply genotyping of polymorphic alleles encoding drug-metabolizing enzymes to the identification of an individual's drug responsiveness phenotype. This knowledge, when applied to dosing or drug selection, can avoid adverse reactions or therapeutic failure and thus enhance therapeutic or prophylactic efficiency when treating a subject with a ORFX modulator, such as a modulator identified by one of the exemplary screening assays described herein.

Monitoring Clinical Efficacy

Monitoring the influence of agents (*e.g.*, drugs, compounds) on the expression or activity of ORFX (*e.g.*, the ability to modulate aberrant cell proliferation and/or differentiation) can be applied in basic drug screening and in clinical trials. For example, the effectiveness of an agent
5 determined by a screening assay as described herein to increase ORFX gene expression, protein levels, or upregulate ORFX activity, can be monitored in clinical trials of subjects exhibiting decreased ORFX gene expression, protein levels, or downregulated ORFX activity.

Alternatively, the effectiveness of an agent determined by a screening assay to decrease ORFX gene expression, protein levels, or downregulate ORFX activity, can be monitored in clinical
10 trials of subjects exhibiting increased ORFX gene expression, protein levels, or upregulated ORFX activity. In such clinical trials, the expression or activity of ORFX and, preferably, other genes that have been implicated in, for example, a proliferative or neurological disorder, can be used as a "read out" or marker of the responsiveness of a particular cell.

For example, genes, including ORFX, that are modulated in cells by treatment with an
15 agent (*e.g.*, compound, drug or small molecule) that modulates ORFX activity (*e.g.*, identified in a screening assay as described herein) can be identified. Thus, to study the effect of agents on cellular proliferation disorders, for example, in a clinical trial, cells can be isolated and RNA prepared and analyzed for the levels of expression of ORFX and other genes implicated in the disorder. The levels of gene expression (*i.e.*, a gene expression pattern) can be quantified by
20 Northern blot analysis or RT-PCR, as described herein, or alternatively by measuring the amount of protein produced, by one of the methods as described herein, or by measuring the levels of activity of ORFX or other genes. In this way, the gene expression pattern can serve as a marker, indicative of the physiological response of the cells to the agent. Accordingly, this response state may be determined before, and at various points during, treatment of the individual with the
25 agent.

In one embodiment, the invention provides a method for monitoring the effectiveness of treatment of a subject with an agent (*e.g.*, an agonist, antagonist, protein, peptide, nucleic acid, peptidomimetic, small molecule, or other drug candidate identified by the screening assays described herein) comprising the steps of (i) obtaining a pre-administration sample from a
30 subject prior to administration of the agent; (ii) detecting the level of expression of a ORFX protein, mRNA, or genomic DNA in the preadministration sample; (iii) obtaining one or more

post-administration samples from the subject; (iv) detecting the level of expression or activity of the ORFX protein, mRNA, or genomic DNA in the post-administration samples; (v) comparing the level of expression or activity of the ORFX protein, mRNA, or genomic DNA in the pre-administration sample with the ORFX protein, mRNA, or genomic DNA in the post administration sample or samples; and (vi) altering the administration of the agent to the subject accordingly. For example, increased administration of the agent may be desirable to increase the expression or activity of ORFX to higher levels than detected, *i.e.*, to increase the effectiveness of the agent. Alternatively, decreased administration of the agent may be desirable to decrease expression or activity of ORFX to lower levels than detected, *i.e.*, to decrease the effectiveness of the agent.

Methods of Treatment

The present invention provides for both prophylactic and therapeutic methods of treating a subject at risk of (or susceptible to) a disorder or having a disorder associated with aberrant ORFX expression or activity.

Diseases and disorders that are characterized by increased (relative to a subject not suffering from the disease or disorder) levels or biological activity may be treated with Therapeutics that antagonize (*i.e.*, reduce or inhibit) activity. Therapeutics that antagonize activity may be administered in a therapeutic or prophylactic manner. Therapeutics that may be utilized include, but are not limited to, (i) a ORFX polypeptide, or analogs, derivatives, fragments or homologs thereof; (ii) antibodies to a ORFX peptide; (iii) nucleic acids encoding a ORFX peptide; (iv) administration of antisense nucleic acid and nucleic acids that are "dysfunctional" (*i.e.*, due to a heterologous insertion within the coding sequences of coding sequences to a ORFX peptide) that are utilized to "knockout" endogenous function of a ORFX peptide by homologous recombination (see, *e.g.*, Capecchi, 1989, *Science* 244: 1288-1292); or (v) modulators (*i.e.*, inhibitors, agonists and antagonists, including additional peptide mimetic of the invention or antibodies specific to a peptide of the invention) that alter the interaction between a ORFX peptide and its binding partner.

Diseases and disorders that are characterized by decreased (relative to a subject not suffering from the disease or disorder) levels or biological activity may be treated with Therapeutics that increase (*i.e.*, are agonists to) activity. Therapeutics that upregulate activity may be administered in a therapeutic or prophylactic manner. Therapeutics that may be utilized

include, but are not limited to, a ORFX peptide, or analogs, derivatives, fragments or homologs thereof; or an agonist that increases bioavailability.

Increased or decreased levels can be readily detected by quantifying peptide and/or RNA, by obtaining a patient tissue sample (*e.g.*, from biopsy tissue) and assaying it *in vitro* for RNA or peptide levels, structure and/or activity of the expressed peptides (or mRNAs of a ORFX peptide). Methods that are well-known within the art include, but are not limited to, immunoassays (*e.g.*, by Western blot analysis, immunoprecipitation followed by sodium dodecyl sulfate (SDS) polyacrylamide gel electrophoresis, immunocytochemistry, etc.) and/or hybridization assays to detect expression of mRNAs (*e.g.*, Northern assays, dot blots, *in situ* hybridization, etc.).

In one aspect, the invention provides a method for preventing, in a subject, a disease or condition associated with an aberrant ORFX expression or activity, by administering to the subject an agent that modulates ORFX expression or at least one ORFX activity. Subjects at risk for a disease that is caused or contributed to by aberrant ORFX expression or activity can be identified by, for example, any or a combination of diagnostic or prognostic assays as described herein. Administration of a prophylactic agent can occur prior to the manifestation of symptoms characteristic of the ORFX aberrancy, such that a disease or disorder is prevented or, alternatively, delayed in its progression. Depending on the type of ORFX aberrancy, for example, a ORFX agonist or ORFX antagonist agent can be used for treating the subject. The appropriate agent can be determined based on screening assays described herein.

Another aspect of the invention pertains to methods of modulating ORFX expression or activity for therapeutic purposes. The modulatory method of the invention involves contacting a cell with an agent that modulates one or more of the activities of ORFX protein activity associated with the cell. An agent that modulates ORFX protein activity can be an agent as described herein, such as a nucleic acid or a protein, a naturally-occurring cognate ligand of a ORFX protein, a peptide, a ORFX peptidomimetic, or other small molecule. In one embodiment, the agent stimulates one or more ORFX protein activity. Examples of such stimulatory agents include active ORFX protein and a nucleic acid molecule encoding ORFX that has been introduced into the cell. In another embodiment, the agent inhibits one or more ORFX protein activity. Examples of such inhibitory agents include antisense ORFX nucleic acid molecules and anti-ORFX antibodies. These modulatory methods can be performed *in vitro*

(*e.g.*, by culturing the cell with the agent) or, alternatively, *in vivo* (*e.g.*, by administering the agent to a subject). As such, the present invention provides methods of treating an individual afflicted with a disease or disorder characterized by aberrant expression or activity of a ORFX protein or nucleic acid molecule. In one embodiment, the method involves administering an agent (*e.g.*, an agent identified by a screening assay described herein), or combination of agents that modulates (*e.g.*, upregulates or downregulates) ORFX expression or activity. In another embodiment, the method involves administering a ORFX protein or nucleic acid molecule as therapy to compensate for reduced or aberrant ORFX expression or activity.

Determination of the Biological Effect of a Therapeutic

In various embodiments of the present invention, suitable *in vitro* or *in vivo* assays are utilized to determine the effect of a specific Therapeutic and whether its administration is indicated for treatment of the affected tissue.

In various specific embodiments, *in vitro* assays may be performed with representative cells of the type(s) involved in the patient's disorder, to determine if a given Therapeutic exerts the desired effect upon the cell type(s). Compounds for use in therapy may be tested in suitable animal model systems including, but not limited to rats, mice, chicken, cows, monkeys, rabbits, and the like, prior to testing in human subjects. Similarly, for *in vivo* testing, any of the animal model system known in the art may be used prior to administration to human subjects.

Malignancies

Some ORFX polypeptides are expressed in cancerous cells (*see, e.g.*, Tables 1 and 2). Accordingly, the corresponding ORF protein is involved in the regulation of cell proliferation. Accordingly, Therapeutics of the present invention may be useful in the therapeutic or prophylactic treatment of diseases or disorders that are associated with cell hyperproliferation and/or loss of control of cell proliferation (*e.g.*, cancers, malignancies and tumors). For a review of such hyperproliferation disorders, *see e.g.*, Fishman, *et al.*, 1985. MEDICINE, 2nd ed., J.B. Lippincott Co., Philadelphia, PA.

Therapeutics of the present invention may be assayed by any method known within the art for efficacy in treating or preventing malignancies and related disorders. Such assays include,

but are not limited to, *in vitro* assays utilizing transformed cells or cells derived from the patient's tumor, as well as *in vivo* assays using animal models of cancer or malignancies. Potentially effective Therapeutics are those that, for example, inhibit the proliferation of tumor-derived or transformed cells in culture or cause a regression of tumors in animal models, in comparison to the controls.

In the practice of the present invention, once a malignancy or cancer has been shown to be amenable to treatment by modulating (*i.e.*, inhibiting, antagonizing or agonizing) activity, that cancer or malignancy may subsequently be treated or prevented by the administration of a Therapeutic that serves to modulate protein function.

Premalignant conditions

The Therapeutics of the present invention that are effective in the therapeutic or prophylactic treatment of cancer or malignancies may also be administered for the treatment of pre-malignant conditions and/or to prevent the progression of a pre-malignancy to a neoplastic or malignant state. Such prophylactic or therapeutic use is indicated in conditions known or suspected of preceding progression to neoplasia or cancer, in particular, where non-neoplastic cell growth consisting of hyperplasia, metaplasia or, most particularly, dysplasia has occurred. For a review of such abnormal cell growth see *e.g.*, Robbins & Angell, 1976. BASIC PATHOLOGY, 2nd ed., W.B. Saunders Co., Philadelphia, PA.

Hyperplasia is a form of controlled cell proliferation involving an increase in cell number in a tissue or organ, without significant alteration in its structure or function. For example, it has been demonstrated that endometrial hyperplasia often precedes endometrial cancer. Metaplasia is a form of controlled cell growth in which one type of mature or fully differentiated cell substitutes for another type of mature cell. Metaplasia may occur in epithelial or connective tissue cells. Dysplasia is generally considered a precursor of cancer, and is found mainly in the epithelia. Dysplasia is the most disorderly form of non-neoplastic cell growth, and involves a loss in individual cell uniformity and in the architectural orientation of cells. Dysplasia characteristically occurs where there exists chronic irritation or inflammation, and is often found in the cervix, respiratory passages, oral cavity, and gall bladder.

Alternatively, or in addition to the presence of abnormal cell growth characterized as hyperplasia, metaplasia, or dysplasia, the presence of one or more characteristics of a transformed or malignant phenotype displayed either *in vivo* or *in vitro* within a cell sample derived from a patient, is indicative of the desirability of prophylactic/therapeutic administration of a Therapeutic that possesses the ability to modulate activity of An aforementioned protein. Characteristics of a transformed phenotype include, but are not limited to: (i) morphological changes; (ii) looser substratum attachment; (iii) loss of cell-to-cell contact inhibition; (iv) loss of anchorage dependence; (v) protease release; (vi) increased sugar transport; (vii) decreased serum requirement; (viii) expression of fetal antigens, (ix) disappearance of the 250 kDal cell-surface protein, and the like. See *e.g.*, Richards, *et al.*, 1986. MOLECULAR PATHOLOGY, W.B. Saunders Co., Philadelphia, PA.

In a specific embodiment of the present invention, a patient that exhibits one or more of the following predisposing factors for malignancy is treated by administration of an effective amount of a Therapeutic: (i) a chromosomal translocation associated with a malignancy (*e.g.*, the Philadelphia chromosome (*bcr/abl*) for chronic myelogenous leukemia and t(14;18) for follicular lymphoma, etc.); (ii) familial polyposis or Gardner's syndrome (possible forerunners of colon cancer); (iii) monoclonal gammopathy of undetermined significance (a possible precursor of multiple myeloma) and (iv) a first degree kinship with persons having a cancer or pre-cancerous disease showing a Mendelian (genetic) inheritance pattern (*e.g.*, familial polyposis of the colon, Gardner's syndrome, hereditary exostosis, polyendocrine adenomatosis, Peutz-Jeghers syndrome, neurofibromatosis of Von Recklinghausen, medullary thyroid carcinoma with amyloid production and pheochromocytoma, retinoblastoma, carotid body tumor, cutaneous melanocarcinoma, intraocular melanocarcinoma, xeroderma pigmentosum, ataxia telangiectasia, Chediak-Higashi syndrome, albinism, Fanconi's aplastic anemia and Bloom's syndrome).

In another embodiment, a Therapeutic of the present invention is administered to a human patient to prevent the progression to breast, colon, lung, pancreatic, or uterine cancer, or melanoma or sarcoma.

Hyperproliferative and dysproliferative disorders

In one embodiment of the present invention, a Therapeutic is administered in the therapeutic or prophylactic treatment of hyperproliferative or benign dysproliferative disorders. The efficacy in treating or preventing hyperproliferative diseases or disorders of a Therapeutic of the present invention may be assayed by any method known within the art. Such assays include *in vitro* cell proliferation assays, *in vitro* or *in vivo* assays using animal models of hyperproliferative diseases or disorders, or the like. Potentially effective Therapeutics may, for example, promote cell proliferation in culture or cause growth or cell proliferation in animal models in comparison to controls.

Specific embodiments of the present invention are directed to the treatment or prevention of cirrhosis of the liver (a condition in which scarring has overtaken normal liver regeneration processes); treatment of keloid (hypertrophic scar) formation causing disfiguring of the skin in which the scarring process interferes with normal renewal; psoriasis (a common skin condition characterized by excessive proliferation of the skin and delay in proper cell fate determination); benign tumors; fibrocystic conditions and tissue hypertrophy (*e.g.*, benign prostatic hypertrophy).

Neurodegenerative disorders

Some ORFX proteins are found in cell types have been implicated in the deregulation of cellular maturation and apoptosis, which are both characteristic of neurodegenerative disease. Accordingly, Therapeutics of the invention, particularly but not limited to those that modulate (or supply) activity of an aforementioned protein, may be effective in treating or preventing neurodegenerative disease. Therapeutics of the present invention that modulate the activity of an aforementioned protein involved in neurodegenerative disorders can be assayed by any method known in the art for efficacy in treating or preventing such neurodegenerative diseases and disorders. Such assays include *in vitro* assays for regulated cell maturation or inhibition of apoptosis or *in vivo* assays using animal models of neurodegenerative diseases or disorders, or any of the assays described below. Potentially effective Therapeutics, for example but not by way of limitation, promote regulated cell maturation and prevent cell apoptosis in culture, or reduce neurodegeneration in animal models in comparison to controls.

Once a neurodegenerative disease or disorder has been shown to be amenable to treatment by modulation activity, that neurodegenerative disease or disorder can be treated or prevented by administration of a Therapeutic that modulates activity. Such diseases include all degenerative disorders involved with aging, especially osteoarthritis and neurodegenerative disorders.

Disorders related to organ transplantation

Some ORFX can be associated with disorders related to organ transplantation, in particular but not limited to organ rejection. Therapeutics of the invention, particularly those that modulate (or supply) activity, may be effective in treating or preventing diseases or disorders related to organ transplantation. Therapeutics of the invention (particularly Therapeutics that modulate the levels or activity of an aforementioned protein) can be assayed by any method known in the art for efficacy in treating or preventing such diseases and disorders related to organ transplantation. Such assays include *in vitro* assays for using cell culture models as described below, or *in vivo* assays using animal models of diseases and disorders related to organ transplantation, see *e.g.*, below. Potentially effective Therapeutics, for example but not by way of limitation, reduce immune rejection responses in animal models in comparison to controls.

Accordingly, once diseases and disorders related to organ transplantation are shown to be amenable to treatment by modulation of activity, such diseases or disorders can be treated or prevented by administration of a Therapeutic that modulates activity.

Cardiovascular Disease

GENX has been implicated in cardiovascular disorders, including in atherosclerotic plaque formation. Diseases such as cardiovascular disease, including cerebral thrombosis or hemorrhage, ischemic heart or renal disease, peripheral vascular disease, or thrombosis of other major vessel, and other diseases, including diabetes mellitus, hypertension, hypothyroidism, cholesterol ester storage disease, systemic lupus erythematosus, homocysteinemia, and familial protein or lipid processing diseases, and the like, are either directly or indirectly associated with atherosclerosis. Accordingly, Therapeutics of the invention, particularly those that modulate (or supply) activity or formation may be effective in treating or preventing

atherosclerosis-associated diseases or disorders. Therapeutics of the invention (particularly Therapeutics that modulate the levels or activity) can be assayed by any method known in the art, including those described below, for efficacy in treating or preventing such diseases and disorders.

5 A vast array of animal and cell culture models exist for processes involved in atherosclerosis. A limited and non-exclusive list of animal models includes knockout mice for premature atherosclerosis (Kurabayashi and Yazaki, 1996, *Int. Angiol.* 15: 187-194), transgenic mouse models of atherosclerosis (Kappel *et al.*, 1994, *FASEB J.* 8: 583-592), antisense oligonucleotide treatment of animal models (Callow, 1995, *Curr. Opin. Cardiol.* 10: 569-576),
10 transgenic rabbit models for atherosclerosis (Taylor, 1997, *Ann. N.Y. Acad. Sci.* 811: 146-152), hypercholesterolemic animal models (Rosenfeld, 1996, *Diabetes Res. Clin. Pract.* 30 Suppl.: 1-11), hyperlipidemic mice (Paigen *et al.*, 1994, *Curr. Opin. Lipidol.* 5: 258-264), and inhibition of lipoxygenase in animals (Sigal *et al.*, 1994, *Ann. N.Y. Acad. Sci.* 714: 211-224). In addition, *in vitro* cell models include but are not limited to monocytes exposed to low density lipoprotein
15 (Frostegard *et al.*, 1996, *Atherosclerosis* 121: 93-103), cloned vascular smooth muscle cells (Suttles *et al.*, 1995, *Exp. Cell Res.* 218: 331-338), endothelial cell-derived chemoattractant exposed T cells (Katz *et al.*, 1994, *J. Leukoc. Biol.* 55: 567-573), cultured human aortic endothelial cells (Farber *et al.*, 1992, *Am. J. Physiol.* 262: H1088-1085), and foam cell cultures (Libby *et al.*, 1996, *Curr Opin Lipidol* 7: 330-335). Potentially effective Therapeutics, for
20 example but not by way of limitation, reduce foam cell formation in cell culture models, or reduce atherosclerotic plaque formation in hypercholesterolemic mouse models of atherosclerosis in comparison to controls.

 Accordingly, once an atherosclerosis-associated disease or disorder has been shown to be amenable to treatment by modulation of activity or formation, that disease or disorder can be
25 treated or prevented by administration of a Therapeutic that modulates activity.

Cytokine and Cell Proliferation/Differentiation Activity

 A GENX protein of the present invention may exhibit cytokine, cell proliferation (either inducing or inhibiting) or cell differentiation (either inducing or inhibiting) activity or may induce production of other cytokines in certain cell populations. Many protein factors discovered

to date, including all known cytokines, have exhibited activity in one or more factor dependent cell proliferation assays, and hence the assays serve as a convenient confirmation of cytokine activity. The activity of a protein of the present invention is evidenced by any one of a number of routine factor dependent cell proliferation assays for cell lines including, without limitation, 32D, DA2, DA1G, T10, B9, B9/11, BaF3, MC9/G, M+ (preB M+), 2E8, RB5, DA1, 123, T1165, HT2, CTLL2, TF-1, Mo7e and CMK.

The activity of a protein of the invention may, among other means, be measured by the following methods: Assays for T-cell or thymocyte proliferation include without limitation those described in: CURRENT PROTOCOLS IN IMMUNOLOGY, Ed by Coligan *et al.*, Greene Publishing Associates and Wiley-Interscience (Chapter 3 and Chapter 7); Takai *et al.*, *J Immunol* 137:3494-3500, 1986; Bertagnoli *et al.*, *J Immunol* 145:1706-1712, 1990; Bertagnoli *et al.*, *Cell Immunol* 133:327-341, 1991; Bertagnoli, *et al.*, *J Immunol* 149:3778-3783, 1992; Bowman *et al.*, *J Immunol* 152:1756-1761, 1994.

Assays for cytokine production and/or proliferation of spleen cells, lymph node cells or thymocytes include, without limitation, those described by Kruisbeek and Shevach, In: CURRENT PROTOCOLS IN IMMUNOLOGY. Coligan *et al.*, eds. Vol 1, pp. 3.12.1-14, John Wiley and Sons, Toronto 1994; and by Schreiber, In: CURRENT PROTOCOLS IN IMMUNOLOGY. Coligan eds. Vol 1 pp. 6.8.1-8, John Wiley and Sons, Toronto 1994.

Assays for proliferation and differentiation of hematopoietic and lymphopoietic cells include, without limitation, those described by Bottomly *et al.*, In: CURRENT PROTOCOLS IN IMMUNOLOGY. Coligan *et al.*, eds. Vol 1 pp. 6.3.1-6.3.12, John Wiley and Sons, Toronto 1991; deVries *et al.*, *J Exp Med* 173:1205-1211, 1991; Moreau *et al.*, *Nature* 336:690-692, 1988; Greenberger *et al.*, *Proc Natl Acad Sci U.S.A.* 80:2931-2938, 1983; Nordan, In: CURRENT PROTOCOLS IN IMMUNOLOGY. Coligan *et al.*, eds. Vol 1 pp. 6.6.1-5, John Wiley and Sons, Toronto 1991; Smith *et al.*, *Proc Natl Acad Sci U.S.A.* 83:1857-1861, 1986; Measurement of human Interleukin 11-Bennett, *et al.* In: CURRENT PROTOCOLS IN IMMUNOLOGY. Coligan *et al.*, eds. Vol 1 pp. 6.15.1 John Wiley and Sons, Toronto 1991; Ciarletta, *et al.*, In: CURRENT PROTOCOLS IN IMMUNOLOGY. Coligan *et al.*, eds. Vol 1 pp. 6.13.1, John Wiley and Sons, Toronto 1991.

Assays for T-cell clone responses to antigens (which will identify, among others, proteins that affect APC-T cell interactions as well as direct T-cell effects by measuring proliferation and

cytokine production) include, without limitation, those described In: CURRENT PROTOCOLS IN IMMUNOLOGY. Coligan *et al.*, eds., Greene Publishing Associates and Wiley-Interscience (Chapter 3Chapter 6, Chapter 7); Weinberger *et al.*, *Proc Natl Acad Sci USA* 77:6091-6095, 1980; Weinberger *et al.*, *Eur J Immunol* 11:405-411, 1981; Takai *et al.*, *J Immunol* 137:3494-3500, 1986; Takai *et al.*, *J Immunol* 140:508-512, 1988.

Immune Stimulating or Suppressing Activity

A GENX protein of the present invention may also exhibit immune stimulating or immune suppressing activity, including without limitation the activities for which assays are described herein. A protein may be useful in the treatment of various immune deficiencies and disorders (including severe combined immunodeficiency (SCID)), *e.g.*, in regulating (up or down) growth and proliferation of T and/or B lymphocytes, as well as effecting the cytolytic activity of NK cells and other cell populations. These immune deficiencies may be genetic or be caused by vital (*e.g.*, HIV) as well as bacterial or fungal infections, or may result from autoimmune disorders. More specifically, infectious diseases causes by vital, bacterial, fungal or other infection may be treatable using a protein of the present invention, including infections by HIV, hepatitis viruses, herpesviruses, mycobacteria, Leishmania species., malaria species. and various fungal infections such as candidiasis. Of course, in this regard, a protein of the present invention may also be useful where a boost to the immune system generally may be desirable, *i.e.*, in the treatment of cancer.

Autoimmune disorders which may be treated using a protein of the present invention include, for example, connective tissue disease, multiple sclerosis, systemic lupus erythematosus, rheumatoid arthritis, autoimmune pulmonary inflammation, Guillain-Barre syndrome, autoimmune thyroiditis, insulin dependent diabetes mellitus, myasthenia gravis, graft-versus-host disease and autoimmune inflammatory eye disease. Such a protein of the present invention may also to be useful in the treatment of allergic reactions and conditions, such as asthma (particularly allergic asthma) or other respiratory problems. Other conditions, in which immune suppression is desired (including, for example, organ transplantation), may also be treatable using a protein of the present invention.

Using the proteins of the invention it may also be possible to immune responses, in a number of ways. Down regulation may be in the form of inhibiting or blocking an immune response already in progress or may involve preventing the induction of an immune response.

The functions of activated T cells may be inhibited by suppressing T cell responses or by inducing specific tolerance in T cells, or both. Immunosuppression of T cell responses is generally an active, non-antigen-specific, process which requires continuous exposure of the T cells to the suppressive agent. Tolerance, which involves inducing non-responsiveness or energy in T cells, is distinguishable from immunosuppression in that it is generally antigen-specific and persists after exposure to the tolerizing agent has ceased. Operationally, tolerance can be demonstrated by the lack of a T cell response upon re-exposure to specific antigen in the absence of the tolerizing agent.

Down regulating or preventing one or more antigen functions (including without limitation B lymphocyte antigen functions (such as, for example, B7), *e.g.*, preventing high level lymphokine synthesis by activated T cells, will be useful in situations of tissue, skin and organ transplantation and in graft-versus-host disease (GVHD). For example, blockage of T cell function should result in reduced tissue destruction in tissue transplantation. Typically, in tissue transplants, rejection of the transplant is initiated through its recognition as foreign by T cells, followed by an immune reaction that destroys the transplant. The administration of a molecule which inhibits or blocks interaction of a B7 lymphocyte antigen with its natural ligand(s) on immune cells (such as a soluble, monomeric form of a peptide having B7-2 activity alone or in conjunction with a monomeric form of a peptide having an activity of another B lymphocyte antigen (*e.g.*, B7-1, B7-3) or blocking antibody), prior to transplantation can lead to the binding of the molecule to the natural ligand(s) on the immune cells without transmitting the corresponding costimulatory signal. Blocking B lymphocyte antigen function in this matter prevents cytokine synthesis by immune cells, such as T cells, and thus acts as an immunosuppressant. Moreover, the lack of costimulation may also be sufficient to energize the T cells, thereby inducing tolerance in a subject. Induction of long-term tolerance by B lymphocyte antigen-blocking reagents may avoid the necessity of repeated administration of these blocking reagents. To achieve sufficient immunosuppression or tolerance in a subject, it may also be necessary to block the function of B lymphocyte antigens.

The efficacy of particular blocking reagents in preventing organ transplant rejection or GVHD can be assessed using animal models that are predictive of efficacy in humans. Examples of appropriate systems which can be used include allogeneic cardiac grafts in rats and xenogeneic pancreatic islet cell grafts in mice, both of which have been used to examine the

immunosuppressive effects of CTLA4Ig fusion proteins *in vivo* as described in Lenschow *et al.*, Science 257:789-792 (1992) and Turka *et al.*, Proc Natl Acad Sci USA, 89:11102-11105 (1992). In addition, murine models of GVHD (see Paul ed., FUNDAMENTAL IMMUNOLOGY, Raven Press, New York, 1989, pp. 846-847) can be used to determine the effect of blocking B lymphocyte antigen function *in vivo* on the development of that disease.

Blocking antigen function may also be therapeutically useful for treating autoimmune diseases. Many autoimmune disorders are the result of inappropriate activation of T cells that are reactive against self tissue and which promote the production of cytokines and auto-antibodies involved in the pathology of the diseases. Preventing the activation of autoreactive T cells may reduce or eliminate disease symptoms. Administration of reagents which block costimulation of T cells by disrupting receptor:ligand interactions of B lymphocyte antigens can be used to inhibit T cell activation and prevent production of auto-antibodies or T cell-derived cytokines which may be involved in the disease process. Additionally, blocking reagents may induce antigen-specific tolerance of autoreactive T cells which could lead to long-term relief from the disease. The efficacy of blocking reagents in preventing or alleviating autoimmune disorders can be determined using a number of well-characterized animal models of human autoimmune diseases. Examples include murine experimental autoimmune encephalitis, systemic lupus erythematosus in MRL/lpr/lpr mice or NZB hybrid mice, murine autoimmune collagen arthritis, diabetes mellitus in NOD mice and BB rats, and murine experimental myasthenia gravis (see Paul ed., FUNDAMENTAL IMMUNOLOGY, Raven Press, New York, 1989, pp. 840-856).

Upregulation of an antigen function (preferably a B lymphocyte antigen function), as a means of up regulating immune responses, may also be useful in therapy. Upregulation of immune responses may be in the form of enhancing an existing immune response or eliciting an initial immune response. For example, enhancing an immune response through stimulating B lymphocyte antigen function may be useful in cases of viral infection. In addition, systemic vital diseases such as influenza, the common cold, and encephalitis might be alleviated by the administration of stimulatory forms of B lymphocyte antigens systemically.

Alternatively, anti-viral immune responses may be enhanced in an infected patient by removing T cells from the patient, costimulating the T cells *in vitro* with viral antigen-pulsed APCs either expressing a peptide of the present invention or together with a stimulatory form of a soluble peptide of the present invention and reintroducing the *in vitro* activated T cells into the

patient. Another method of enhancing anti-vital immune responses would be to isolate infected cells from a patient, transfect them with a nucleic acid encoding a protein of the present invention as described herein such that the cells express all or a portion of the protein on their surface, and reintroduce the transfected cells into the patient. The infected cells would now be
5 capable of delivering a costimulatory signal to, and thereby activate, T cells *in vivo*.

In another application, up regulation or enhancement of antigen function (preferably B lymphocyte antigen function) may be useful in the induction of tumor immunity. Tumor cells (e.g., sarcoma, melanoma, lymphoma, leukemia, neuroblastoma, carcinoma) transfected with a nucleic acid encoding at least one peptide of the present invention can be administered to a
10 subject to overcome tumor-specific tolerance in the subject. If desired, the tumor cell can be transfected to express a combination of peptides. For example, tumor cells obtained from a patient can be transfected *ex vivo* with an expression vector directing the expression of a peptide having B7-2-like activity alone, or in conjunction with a peptide having B7-1-like activity and/or B7-3-like activity. The transfected tumor cells are returned to the patient to result in expression
15 of the peptides on the surface of the transfected cell. Alternatively, gene therapy techniques can be used to target a tumor cell for transfection *in vivo*.

The presence of the peptide of the present invention having the activity of a B lymphocyte antigen(s) on the surface of the tumor cell provides the necessary costimulation signal to T cells to induce a T cell mediated immune response against the transfected tumor cells.
20 In addition, tumor cells which lack MHC class I or MHC class II molecules, or which fail to reexpress sufficient amounts of MHC class I or MHC class II molecules, can be transfected with nucleic acid encoding all or a portion of (e.g., a cytoplasmic-domain truncated portion) of an MHC class I α chain protein and β_2 microglobulin protein or an MHC class II α chain protein and an MHC class II β chain protein to thereby express MHC class I or MHC class II proteins on
25 the cell surface. Expression of the appropriate class I or class II MHC in conjunction with a peptide having the activity of a B lymphocyte antigen (e.g., B7-1, B7-2, B7-3) induces a T cell mediated immune response against the transfected tumor cell. Optionally, a gene encoding an antisense construct which blocks expression of an MHC class II associated protein, such as the invariant chain, can also be cotransfected with a DNA encoding a peptide having the activity of a
30 B lymphocyte antigen to promote presentation of tumor associated antigens and induce tumor

specific immunity. Thus, the induction of a T cell mediated immune response in a human subject may be sufficient to overcome tumor-specific tolerance in the subject.

The activity of a protein of the invention may, among other means, be measured by the following methods: Suitable assays for thymocyte or splenocyte cytotoxicity include, without
5 limitation, those described In: CURRENT PROTOCOLS IN IMMUNOLOGY. Coligan *et al.*, eds. Greene Publishing Associates and Wiley-Interscience (Chapter 3, Chapter 7); Herrmann *et al.*, *Proc Natl Acad Sci USA* 78:2488-2492, 1981; Herrmann *et al.*, *J Immunol* 128:1968-1974, 1982; Handa *et al.*, *J Immunol* 135:1564-1572, 1985; Takai *et al.*, *J Immunol* 137:3494-3500, 1986; Takai *et al.*, *J Immunol* 140:508-512, 1988; Herrmann *et al.*, *Proc Natl Acad Sci USA* 78:2488-2492, 1981;
10 Herrmann *et al.*, *J Immunol* 128:1968-1974, 1982; Handa *et al.*, *J Immunol* 135:1564-1572, 1985; Takai *et al.*, *J Immunol* 137:3494-3500, 1986; Bowman *et al.*, *J Virology* 61:1992-1998; Takai *et al.*, *J Immunol* 140:508-512, 1988; Bertagnolli *et al.*, *Cell Immunol* 133:327-341, 1991; Brown *et al.*, *J Immunol* 153:3079-3092, 1994.

Assays for T-cell-dependent immunoglobulin responses and isotype switching (which
15 will identify, among others, proteins that modulate T-cell dependent antibody responses and that affect Th1/Th2 profiles) include, without limitation, those described in: Maliszewski, *J Immunol* 144:3028-3033, 1990; and Mond and Brunswick In: CURRENT PROTOCOLS IN IMMUNOLOGY. Coligan *et al.*, (eds.) Vol 1 pp. 3.8.1-3.8.16, John Wiley and Sons, Toronto 1994.

Mixed lymphocyte reaction (MLR) assays (which will identify, among others, proteins
20 that generate predominantly Th1 and CTL responses) include, without limitation, those described In: CURRENT PROTOCOLS IN IMMUNOLOGY. Coligan *et al.*, eds. Greene Publishing Associates and Wiley-Interscience (Chapter 3, Chapter 7); Takai *et al.*, *J Immunol* 137:3494-3500, 1986; Takai *et al.*, *J Immunol* 140:508-512, 1988; Bertagnolli *et al.*, *J Immunol* 149:3778-3783, 1992.

Dendritic cell-dependent assays (which will identify, among others, proteins expressed by
25 dendritic cells that activate naive T-cells) include, without limitation, those described in: Guery *et al.*, *J Immunol* 134:536-544, 1995; Inaba *et al.*, *J Exp Med* 173:549-559, 1991; Macatonia *et al.*, *J Immunol* 154:5071-5079, 1995; Porgador *et al.*, *J Exp Med* 182:255-260, 1995; Nair *et al.*, *J Virol* 67:4062-4069, 1993; Huang *et al.*, *Science* 264:961-965, 1994; Macatonia *et al.*, *J Exp Med* 169:1255-1264, 1989; Bhardwaj *et al.*, *J Clin Investig* 94:797-807, 1994; and Inaba *et al.*, *J*
30 *Exp Med* 172:631-640, 1990.

Assays for lymphocyte survival/apoptosis (which will identify, among others, proteins that prevent apoptosis after superantigen induction and proteins that regulate lymphocyte homeostasis) include, without limitation, those described in: Darzynkiewicz *et al.*, *Cytometry* 13:795-808, 1992; Gorczyca *et al.*, *Leukemia* 7:659-670, 1993; Gorczyca *et al.*, *Cancer Res* 53:1945-1951, 1993; Itoh *et al.*, *Cell* 66:233-243, 1991; Zacharchuk, *J Immunol* 145:4037-4045, 1990; Zamai *et al.*, *Cytometry* 14:891-897, 1993; Gorczyca *et al.*, *Internat J Oncol* 1:639-648, 1992.

Assays for proteins that influence early steps of T-cell commitment and development include, without limitation, those described in: Antica *et al.*, *Blood* 84:111-117, 1994; Fine *et al.*, *Cell Immunol* 155: 111-122, 1994; Galy *et al.*, *Blood* 85:2770-2778, 1995; Toki *et al.*, *Proc Nat Acad Sci USA* 88:7548-7551, 1991.

Hematopoiesis Regulating Activity

A GENX protein of the present invention may be useful in regulation of hematopoiesis and, consequently, in the treatment of myeloid or lymphoid cell deficiencies. Even marginal biological activity in support of colony forming cells or of factor-dependent cell lines indicates involvement in regulating hematopoiesis, *e.g.* in supporting the growth and proliferation of erythroid progenitor cells alone or in combination with other cytokines, thereby indicating utility, for example, in treating various anemias or for use in conjunction with irradiation/chemotherapy to stimulate the production of erythroid precursors and/or erythroid cells; in supporting the growth and proliferation of myeloid cells such as granulocytes and monocytes/macrophages (*i.e.*, traditional CSF activity) useful, for example, in conjunction with chemotherapy to prevent or treat consequent myelo-suppression; in supporting the growth and proliferation of megakaryocytes and consequently of platelets thereby allowing prevention or treatment of various platelet disorders such as thrombocytopenia, and generally for use in place of or complimentary to platelet transfusions; and/or in supporting the growth and proliferation of hematopoietic stem cells which are capable of maturing to any and all of the above-mentioned hematopoietic cells and therefore find therapeutic utility in various stem cell disorders (such as those usually treated with transplantation, including, without limitation, aplastic anemia and paroxysmal nocturnal hemoglobinuria), as well as in repopulating the stem cell compartment post irradiation/chemotherapy, either in-vivo or ex-vivo (*i.e.*, in conjunction with bone marrow

transplantation or with peripheral progenitor cell transplantation (homologous or heterologous)) as normal cells or genetically manipulated for gene therapy.

The activity of a protein of the invention may, among other means, be measured by the following methods:

5 Suitable assays for proliferation and differentiation of various hematopoietic lines are cited above.

Assays for embryonic stem cell differentiation (which will identify, among others, proteins that influence embryonic differentiation hematopoiesis) include, without limitation, those described in: Johansson *et al.* *Cellular Biology* 15:141-151, 1995; Keller *et al.*, *Mol. Cell.*
10 *Biol.* 13:473-486, 1993; McClanahan *et al.*, *Blood* 81:2903-2915, 1993.

Assays for stem cell survival and differentiation (which will identify, among others, proteins that regulate lympho-hematopoiesis) include, without limitation, those described in: Methylcellulose colony forming assays, Freshney, In: CULTURE OF HEMATOPOIETIC CELLS. Freshney, *et al.* (eds.) Vol pp. 265-268, Wiley-Liss, Inc., New York, N.Y. 1994; Hirayama *et al.*,
15 *Proc Natl Acad Sci USA* 89:5907-5911, 1992; McNiece and Briddeli, In: CULTURE OF HEMATOPOIETIC CELLS. Freshney, *et al.* (eds.) Vol pp. 23-39, Wiley-Liss, Inc., New York, N.Y. 1994; Neben *et al.*, *Exp Hematol* 22:353-359, 1994; Ploemacher, In: CULTURE OF HEMATOPOIETIC CELLS. Freshney, *et al.* eds. Vol pp. 1-21, Wiley-Liss, Inc., New York, N.Y. 1994; Spoonceret *et al.*, In: CULTURE OF HEMATOPOIETIC CELLS. Freshney, *et al.*, (eds.) Vol pp.
20 163-179, Wiley-Liss, Inc., New York, N.Y. 1994; Sutherland, In: CULTURE OF HEMATOPOIETIC CELLS. Freshney, *et al.*, (eds.) Vol pp. 139-162, Wiley-Liss, Inc., New York, N.Y. 1994.

Tissue Growth Activity

A GENX protein of the present invention also may have utility in compositions used for bone, cartilage, tendon, ligament and/or nerve tissue growth or regeneration, as well as for
25 wound healing and tissue repair and replacement, and in the treatment of burns, incisions and ulcers.

A protein of the present invention, which induces cartilage and/or bone growth in circumstances where bone is not normally formed, has application in the healing of bone fractures and cartilage damage or defects in humans and other animals. Such a preparation
30 employing a protein of the invention may have prophylactic use in closed as well as open fracture reduction and also in the improved fixation of artificial joints. De novo bone formation

induced by an osteogenic agent contributes to the repair of congenital, trauma induced, or oncologic resection induced craniofacial defects, and also is useful in cosmetic plastic surgery.

A protein of this invention may also be used in the treatment of periodontal disease, and in other tooth repair processes. Such agents may provide an environment to attract bone-forming cells, stimulate growth of bone-forming cells or induce differentiation of progenitors of bone-forming cells. A protein of the invention may also be useful in the treatment of osteoporosis or osteoarthritis, such as through stimulation of bone and/or cartilage repair or by blocking inflammation or processes of tissue destruction (collagenase activity, osteoclast activity, etc.) mediated by inflammatory processes.

Another category of tissue regeneration activity that may be attributable to the protein of the present invention is tendon/ligament formation. A protein of the present invention, which induces tendon/ligament-like tissue or other tissue formation in circumstances where such tissue is not normally formed, has application in the healing of tendon or ligament tears, deformities and other tendon or ligament defects in humans and other animals. Such a preparation employing a tendon/ligament-like tissue inducing protein may have prophylactic use in preventing damage to tendon or ligament tissue, as well as use in the improved fixation of tendon or ligament to bone or other tissues, and in repairing defects to tendon or ligament tissue. De novo tendon/ligament-like tissue formation induced by a composition of the present invention contributes to the repair of congenital, trauma induced, or other tendon or ligament defects of other origin, and is also useful in cosmetic plastic surgery for attachment or repair of tendons or ligaments. The compositions of the present invention may provide an environment to attract tendon- or ligament-forming cells, stimulate growth of tendon- or ligament-forming cells, induce differentiation of progenitors of tendon- or ligament-forming cells, or induce growth of tendon/ligament cells or progenitors *ex vivo* for return *in vivo* to effect tissue repair. The compositions of the invention may also be useful in the treatment of tendonitis, carpal tunnel syndrome and other tendon or ligament defects. The compositions may also include an appropriate matrix and/or sequestering agent as a carrier as is well known in the art.

The protein of the present invention may also be useful for proliferation of neural cells and for regeneration of nerve and brain tissue, *i.e.* for the treatment of central and peripheral nervous system diseases and neuropathies, as well as mechanical and traumatic disorders, which involve degeneration, death or trauma to neural cells or nerve tissue. More specifically, a protein

may be used in the treatment of diseases of the peripheral nervous system, such as peripheral nerve injuries, peripheral neuropathy and localized neuropathies, and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager syndrome. Further conditions which may be treated in accordance with the present invention include mechanical and traumatic disorders, such as spinal cord disorders, head trauma and cerebrovascular diseases such as stroke. Peripheral neuropathies resulting from chemotherapy or other medical therapies may also be treatable using a protein of the invention.

Proteins of the invention may also be useful to promote better or faster closure of non-healing wounds, including without limitation pressure ulcers, ulcers associated with vascular insufficiency, surgical and traumatic wounds, and the like.

It is expected that a protein of the present invention may also exhibit activity for generation or regeneration of other tissues, such as organs (including, for example, pancreas, liver, intestine, kidney, skin, endothelium), muscle (smooth, skeletal or cardiac) and vascular (including vascular endothelium) tissue, or for promoting the growth of cells comprising such tissues. Part of the desired effects may be by inhibition or modulation of fibrotic scarring to allow normal tissue to regenerate. A protein of the invention may also exhibit angiogenic activity.

A protein of the present invention may also be useful for gut protection or regeneration and treatment of lung or liver fibrosis, reperfusion injury in various tissues, and conditions resulting from systemic cytokine damage.

A protein of the present invention may also be useful for promoting or inhibiting differentiation of tissues described above from precursor tissues or cells; or for inhibiting the growth of tissues described above.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assays for tissue generation activity include, without limitation, those described in: International Patent Publication No. WO95/16035 (bone, cartilage, tendon); International Patent Publication No. WO95/05846 (nerve, neuronal); International Patent Publication No. WO91/07491 (skin, endothelium).

Assays for wound healing activity include, without limitation, those described in: Winter, EPIDERMAL WOUND HEALING, pp. 71-112 (Maibach and Rovee, eds.), Year Book Medical

Publishers, Inc., Chicago, as modified by Eaglstein and Menz, *J. Invest. Dermatol* 71:382-84 (1978).

Activin/Inhibin Activity

5 A GENX protein of the present invention may also exhibit activin- or inhibin-related activities. Inhibins are characterized by their ability to inhibit the release of follicle stimulating hormone (FSH), while activins are characterized by their ability to stimulate the release of follicle stimulating hormone (FSH). Thus, a protein of the present invention, alone or in heterodimers with a member of the inhibin family, may be useful as a contraceptive based on the ability of inhibins to decrease fertility in female mammals and decrease spermatogenesis in
10 male mammals. Administration of sufficient amounts of other inhibins can induce infertility in these mammals. Alternatively, the protein of the invention, as a homodimer or as a heterodimer with other protein subunits of the inhibin-b group, may be useful as a fertility inducing therapeutic, based upon the ability of activin molecules in stimulating FSH release from cells of the anterior pituitary. See, for example, U.S. Pat. No. 4,798,885. A protein of the invention may
15 also be useful for advancement of the onset of fertility in sexually immature mammals, so as to increase the lifetime reproductive performance of domestic animals such as cows, sheep and pigs.

The activity of a protein of the invention may, among other means, be measured by the following methods:

20 Assays for activin/inhibin activity include, without limitation, those described in: Vale *et al.*, *Endocrinology* 91:562-572, 1972; Ling *et al.*, *Nature* 321:779-782, 1986; Vale *et al.*, *Nature* 321:776-779, 1986; Mason *et al.*, *Nature* 318:659-663, 1985; Forage *et al.*, *Proc Natl Acad Sci USA* 83:3091-3095, 1986.

Chemotactic/Chemokinetic Activity

25 A protein of the present invention may have chemotactic or chemokinetic activity (*e.g.*, act as a chemokine) for mammalian cells, including, for example, monocytes, fibroblasts, neutrophils, T-cells, mast cells, eosinophils, epithelial and/or endothelial cells. Chemotactic and chemokinetic proteins can be used to mobilize or attract a desired cell population to a desired site of action. Chemotactic or chemokinetic proteins provide particular advantages in treatment of
30 wounds and other trauma to tissues, as well as in treatment of localized infections. For example,

attraction of lymphocytes, monocytes or neutrophils to tumors or sites of infection may result in improved immune responses against the tumor or infecting agent.

A protein or peptide has chemotactic activity for a particular cell population if it can stimulate, directly or indirectly, the directed orientation or movement of such cell population.

5 Preferably, the protein or peptide has the ability to directly stimulate directed movement of cells. Whether a particular protein has chemotactic activity for a population of cells can be readily determined by employing such protein or peptide in any known assay for cell chemotaxis.

The activity of a protein of the invention may, among other means, be measured by following methods:

10 Assays for chemotactic activity (which will identify proteins that induce or prevent chemotaxis) consist of assays that measure the ability of a protein to induce the migration of cells across a membrane as well as the ability of a protein to induce the adhesion of one cell population to another cell population. Suitable assays for movement and adhesion include, without limitation, those described in: CURRENT PROTOCOLS IN IMMUNOLOGY, Coligan *et al.*,
15 eds. (Chapter 6.12, MEASUREMENT OF ALPHA AND BETA CHEMOKINES 6.12.1-6.12.28); Taub *et al. J Clin Invest* 95:1370-1376, 1995; Lind *et al. APMIS* 103:140-146, 1995; Muller *et al., Eur J Immunol* 25: 1744-1748; Gruber *et al. J Immunol* 152:5860-5867, 1994; Johnston *et al., J Immunol* 153: 1762-1768, 1994.

Hemostatic and Thrombolytic Activity

20 A protein of the invention may also exhibit hemostatic or thrombolytic activity. As a result, such a protein is expected to be useful in treatment of various coagulation disorders (including hereditary disorders, such as hemophilias) or to enhance coagulation and other hemostatic events in treating wounds resulting from trauma, surgery or other causes. A protein of the invention may also be useful for dissolving or inhibiting formation of thromboses and for
25 treatment and prevention of conditions resulting therefrom (such as, for example, infarction of cardiac and central nervous system vessels (*e.g.*, stroke).

The activity of a protein of the invention may, among other means, be measured by the following methods:

30 Assay for hemostatic and thrombolytic activity include, without limitation, those described in: Linet *et al., J. Clin. Pharmacol.* 26:131-140, 1986; Burdick *et al., Thrombosis Res.*

45:413-419, 1987; Humphrey *et al.*, *Fibrinolysis* 5:71-79 (1991); Schaub, *Prostaglandins* 35:467-474, 1988.

Receptor/Ligand Activity

A protein of the present invention may also demonstrate activity as receptors, receptor
5 ligands or inhibitors or agonists of receptor/ligand interactions. Examples of such receptors and
ligands include, without limitation, cytokine receptors and their ligands, receptor kinases and
their ligands, receptor phosphatases and their ligands, receptors involved in cell—cell
interactions and their ligands (including without limitation, cellular adhesion molecules (such as
selectins, integrins and their ligands) and receptor/ligand pairs involved in antigen presentation,
10 antigen recognition and development of cellular and humoral immune responses). Receptors and
ligands are also useful for screening of potential peptide or small molecule inhibitors of the
relevant receptor/ligand interaction. A protein of the present invention (including, without
limitation, fragments of receptors and ligands) may themselves be useful as inhibitors of
receptor/ligand interactions.

15 The activity of a protein of the invention may, among other means, be measured by the
following methods:

Suitable assays for receptor-ligand activity include without limitation those described in:
CURRENT PROTOCOLS IN IMMUNOLOGY, Ed by Coligan, *et al.*, Greene Publishing Associates and
Wiley-Interscience (Chapter 7.28, Measurement of Cellular Adhesion under static conditions
20 7.28.1-7.28.22), Takai *et al.*, *Proc Natl Acad Sci USA* 84:6864-6868, 1987; Bierer *et al.*, *J. Exp.*
Med. 168:1145-1156, 1988; Rosenstein *et al.*, *J. Exp. Med.* 169:149-160 1989; Stoltenborg *et al.*,
J Immunol Methods 175:59-68, 1994; Stitt *et al.*, *Cell* 80:661-670, 1995.

Anti-Inflammatory Activity

Proteins of the present invention may also exhibit anti-inflammatory activity. The
25 anti-inflammatory activity may be achieved by providing a stimulus to cells involved in the
inflammatory response, by inhibiting or promoting cell—cell interactions (such as, for example,
cell adhesion), by inhibiting or promoting chemotaxis of cells involved in the inflammatory
process, inhibiting or promoting cell extravasation, or by stimulating or suppressing production
of other factors which more directly inhibit or promote an inflammatory response. Proteins
30 exhibiting such activities can be used to treat inflammatory conditions including chronic or acute

conditions), including without limitation inflammation associated with infection (such as septic shock, sepsis or systemic inflammatory response syndrome (SIRS)), ischemia-reperfusion injury, endotoxin lethality, arthritis, complement-mediated hyperacute rejection, nephritis, cytokine or chemokine-induced lung injury, inflammatory bowel disease, Crohn's disease or resulting from over production of cytokines such as TNF or IL-1. Proteins of the invention may also be useful to treat anaphylaxis and hypersensitivity to an antigenic substance or material.

Tumor Inhibition Activity

In addition to the activities described above for immunological treatment or prevention of tumors, a protein of the invention may exhibit other anti-tumor activities. A protein may inhibit tumor growth directly or indirectly (such as, for example, via ADCC). A protein may exhibit its tumor inhibitory activity by acting on tumor tissue or tumor precursor tissue, by inhibiting formation of tissues necessary to support tumor growth (such as, for example, by inhibiting angiogenesis), by causing production of other factors, agents or cell types which inhibit tumor growth, or by suppressing, eliminating or inhibiting factors, agents or cell types which promote tumor growth.

Other Activities

A protein of the invention may also exhibit one or more of the following additional activities or effects: inhibiting the growth, infection or function of, or killing, infectious agents, including, without limitation, bacteria, viruses, fungi and other parasites; effecting (suppressing or enhancing) bodily characteristics, including, without limitation, height, weight, hair color, eye color, skin, fat to lean ratio or other tissue pigmentation, or organ or body part size or shape (such as, for example, breast augmentation or diminution, change in bone form or shape); effecting biorhythms or circadian cycles or rhythms; effecting the fertility of male or female subjects; effecting the metabolism, catabolism, anabolism, processing, utilization, storage or elimination of dietary fat, lipid, protein, carbohydrate, vitamins, minerals, cofactors or other nutritional factors or component(s); effecting behavioral characteristics, including, without limitation, appetite, libido, stress, cognition (including cognitive disorders), depression (including depressive disorders) and violent behaviors; providing analgesic effects or other pain reducing effects; promoting differentiation and growth of embryonic stem cells in lineages other than hematopoietic lineages; hormonal or endocrine activity; in the case of enzymes, correcting

deficiencies of the enzyme and treating deficiency-related diseases; treatment of hyperproliferative disorders (such as, for example, psoriasis); immunoglobulin-like activity (such as, for example, the ability to bind antigens or complement); and the ability to act as an antigen in a vaccine composition to raise an immune response against such protein or another material or entity which is cross-reactive with such protein.

Neural disorders in general include Parkinson's disease, Alzheimer's disease, Huntington's disease, multiple sclerosis, amyotrophic lateral sclerosis (ALS), peripheral neuropathy, tumors of the nervous system, exposure to neurotoxins, acute brain injury, peripheral nerve trauma or injury, and other neuropathies, epilepsy, and/or tremors.

10 EQUIVALENTS

From the foregoing detailed description of the specific embodiments of the invention, it should be apparent that particular novel compositions and methods involving nucleic acids, polypeptides, antibodies, detection and treatment have been described. Although these particular embodiments have been disclosed herein in detail, this has been done by way of example for purposes of illustration only, and is not intended to be limiting with respect to the scope of the appended claims that follow. In particular, it is contemplated by the inventors that various substitutions, alterations, and modifications may be made as a matter of routine for a person of ordinary skill in the art to the invention without departing from the spirit and scope of the invention as defined by the claims. Indeed, various modifications of the invention in addition to those described herein will become apparent to those skilled in the art from the foregoing description and accompanying figures. Such modifications are intended to fall within the scope of the appended claims.

THIS PAGE INTENTIONALLY LEFT BLANK

THIS PAGE INTENTIONALLY LEFT BLANK

THIS PAGE INTENTIONALLY LEFT BLANK

THIS PAGE INTENTIONALLY LEFT BLANK

THIS PAGE INTENTIONALLY LEFT BLANK

THIS PAGE INTENTIONALLY LEFT BLANK

THIS PAGE INTENTIONALLY LEFT BLANK

THIS PAGE INTENTIONALLY LEFT BLANK

Table 1

| ORF# | Internal Identification Number | Protein similarity | Protein domain | Protein Classification | Cells or Tissues in which Gene is Expressed |
|------|--------------------------------|---|---|------------------------|--|
| 1 | 13076366 (1, 2) | Novel Protein sim. GBank gij4691395jemb CAB41562.1 - (AL049727) putative large secreted protein [Streptomyces coelicolor] | | UNCLASSIFIED | 264636 |
| 2 | 80248091 (3, 4) | Novel Protein sim. GBank gij2829506 sp P71559 SUCC_MYCTU - SUCCINYL-COA SYNTHETASE BETA CHAIN (SCS-BETA) | Contains protein domain (PF00549) - CoA-ligases | UNCLASSIFIED | 264907, 264600, 264602, 264762, 264769, 264689, 264638, 264587 |
| 3 | 80415924 (5, 6) | | | UNCLASSIFIED | 264910, 264604, 264634, 264905, 264636, 264691, 264907, 264692, 264629 |
| 4 | 82018837 (7, 8) | | | UNCLASSIFIED | 264908, 264909, 264760, 264628, 264635 |
| 5 | 79970035 (9, 10) | | | UNCLASSIFIED | 22279002, 264563 |
| 6 | 79842462 (11, 12) | | Contains protein domain (PF00127) - Copper binding proteins, plastocyanin/azurin family | UNCLASSIFIED | 264908 |
| 7 | 85515576 (13, 14) | Novel Protein sim. GBank gij4415925 gb AAD20157 - (AC006282) unknown protein [Arabidopsis thaliana] | | UNCLASSIFIED | 20281099, 35696052, 264508, 264509, 264905, 264906, 264907, 264908, 264909, 264511, 265006, 264512, 265009, 264910, 264595, 264596, 264758, 264603, 264604, 264760, 264762, 264683, 264766, 264767, 264689, 35695917, 264690, 264692, 264693, 33657109, 264628, 264629, 35696423, 55811576, 35695855, 264630, 264631, 264632, 264634, 264636, 264637, 264638, 264639, 18108385, 264563, 264564, 264566, 264488 |
| 8 | 56924278 (15, 16) | Novel Protein sim. GBank gij585562 sp Q06458 NIRB_KLEPN - NITRITE REDUCTASE (NAD(P)H) LARGE SUBUNIT | | reductase | 264907 |
| 9 | 79394457 (17, 18) | | | UNCLASSIFIED | 265007, 265019, 263972 |
| 10 | 79556459 (19, 20) | | | UNCLASSIFIED | 264906 |
| 11 | 20414027 (21, 22) | | | | 264605 |
| 12 | 94141210 (23, 24) | Novel Protein sim. GBank gij3878145 emb CAA9887.1 - (Z75543) similar to potassium channel protein [Caenorhabditis elegans] | | misc_channel | 264259, 265007, 83373044 |
| 13 | 20750551 (25, 26) | | | UNCLASSIFIED | 264556, 264557, 264564 |
| 14 | 95105114 (27, 28) | Novel Protein sim. GBank gij2832781 emb CAA12645 - (AJ225805) inward potassium channel alpha subunit [Egeria densa] | Contains protein domain (PF00023) - Ank repeat | potassium_channel | 35696286, 35696052, 264510, 35695917, 264691, 264628, 35696423, 264555, 264558, 264559, 83373044 |
| 15 | 20458307 (29, 30) | Novel Protein sim. GBank gij1710791 sp Q10234 RT05_SCHPO - PROBABLE MITOCHONDRIAL 40S RIBOSOMAL PROTEIN S5 | Contains protein domain (PF00333) - Ribosomal protein S5 | ribosomalprot | 264604 |
| 16 | 20760356 (31, 32) | | | | 264555 |

| | | | | | |
|----|-------------------|---|--|----------------|---|
| 17 | 20282744 (33, 34) | Novel Protein sim. GBank gij1174884[sp]P44391[URE1_HAEN - UREASE ALPHA SUBUNIT (UREA AMIDOHYDROLASE) Novel Protein sim. GBank gij2281102 (AC002333) - SF18 isolbg [Arabidopsis thaliana] | Contains protein domain (PF00449) - Urease | | 264600 |
| 18 | 80246804 (35, 36) | | | | 26331827, 264555, 264557, 264638, 264558 |
| 19 | 80076624 (37, 38) | | | UNCLASSIFIED | 22278996, 264907, 264910, 264600, 264693 |
| 20 | 20724558 (39, 40) | Novel Protein sim. GBank gij2506112[sp]P43672[UUP_ECOLI - ABC TRANSPORTER ATP-BINDING PROTEIN UUP Novel Protein sim. GBank gij1730203[sp]P50442[GATM_RAT - GLYCINE AMINOTRANSFERASE PRECURSOR (L- ARGININE:GLYCINE AMINOTRANSFERASE) (TRANSAMIDINASE) (AT) | transport | | 264602 |
| 21 | 80417554 (41, 42) | | | UNCLASSIFIED | 22278995, 264906, 265008, 265010, 265011, 264602, 264605, 264766, 264688, 21908764, 264691, 18108376, 264636, 18108387, 264486 |
| 22 | 11705858 (43, 44) | | | | 264685 |
| 23 | 80419176 (45, 46) | Novel Protein sim. GBank gij1877329[emb]CAB07077] - (Z92771) fadE25 [Mycobacterium tuberculosis] | Contains protein domain (PF00441) - Acyl-CoA dehydrogenase | dehydrogenase | 264488, 264907, 264909, 264600, 264602, 264603, 264605, 264682, 264766, 3283386, 264636, 264486 |
| 24 | 20291697 (47, 48) | | | | 264600 |
| 25 | 80253774 (49, 50) | | | | 264593 |
| 26 | 80255394 (51, 52) | | | UNCLASSIFIED | 22278996, 56182435, 265018, 264566 |
| 27 | 80235795 (53, 54) | Novel Protein sim. GBank gij4808369[emb]CAB42783.1] - (AL049841) putative 30S ribosomal protein S14 [Sireptomyces coelicolor] | Contains protein domain (PF00253) - Ribosomal protein S14p/S29e | ribosomal prot | 18108370, 35696423, 264635, 264555 |
| 28 | 78483561 (55, 56) | | | UNCLASSIFIED | 264638 |
| 29 | 82448765 (57, 58) | Novel Protein sim. GBank gij3122290[sp]O08333[K6PF_STRCO - 6- PHOSPHOFRUCTOKINASE (PHOSPHOFRUCTOKINASE) (PHOSPHOHEXOKINASE) (ATP-PFK) | Contains protein domain (PF00365) - Phosphofructokinase | kinase | 264601, 264762, 264768, 264769, 264638 |
| 30 | 79189333 (59, 60) | | | UNCLASSIFIED | 264908, 265019, 264687, 21906764, |
| 31 | 19848158 (61, 62) | | | UNCLASSIFIED | 21906766 |
| 32 | 82448495 (63, 64) | Novel Protein sim. GBank gij3560504 (AF027770) - unknown [Mycobacterium smegmatis] | | UNCLASSIFIED | 264534 |
| 33 | 78582628 (65, 66) | Novel Protein sim. GBank gij2129003[pir]G64507 - hypothetical protein MJ1685 - Methanococcus jannaschii | | UNCLASSIFIED | 264905, 264605, 264762, 264766, 264687, 264689 |
| 34 | 87467657 (67, 68) | | | UNCLASSIFIED | 264687 |
| 35 | 95005170 (69, 70) | Novel Protein sim. GBank gij5420387[emb]CAB46679.1] - (AJ243459) proteophosphoglycan [Leishmania major] | | UNCLASSIFIED | 60432289, 264600, 264602, 264760, 18108357, 264768, 265020, 264691 |
| 36 | 18642042 (71, 72) | Novel Protein sim. GBank gij3287739[sp]P73538[BIOB_SYNY3 - BIOTIN SYNTHASE (BIOTIN SYNTHETASE) | | UNCLASSIFIED | 264600, 264687, 264558, 264639 |
| 37 | 20369215 (73, 74) | Novel Protein sim. GBank gij2313134[gb]AAD07126.1] - (AE000527) delta-1-pyrroline-5-carboxylate dehydrogenase [Helicobacter pylori 26695] | synthase | | 264566 |
| | | | dehydrogenase | | 264603 |

| | | | | | | |
|----|---------------------|---|--|--------------|--|--|
| 38 | 20466334 (75, 76) | Novel Protein sim. GBank gij3805970[emb]CAA06231] - (AJ004933) periplasmic nitrate reductase, large subunit [Rhodospseudomonas sp.] | | | 264605 | |
| 39 | 94300715 (77, 78) | Novel Protein sim. GBank gij1929448 (L63543) - endodermis [Xenopus laevis] | Contains protein domain (PF00207) - Alpha-2-macroglobulin family | complement | 264905, 264906, 264907, 66712502, 264908, 264909, 264511, 265009, 264910, 55812038, 264758, 265011, 264782, 264882, 264763, 264784, 264766, 265022, 264693, 264628, 264631, 264634, 264635, 264555, 264638, 18108381, 264558, 18108385, 264482 | |
| 40 | 20635625 (78, 80) | Novel Protein sim. GBank gij854065[emb]CAA58337] - (X83413) U88 [Human herpesvirus 8] | | UNCLASSIFIED | 264592 | |
| 41 | 80023287 (81, 82) | Novel Protein sim. GBank gij3820584 (AF086791) - carbamoylphosphate synthetase large subunit [Zymomonas mobilis] | | UNCLASSIFIED | 264591, 35695917 | |
| 42 | 20724566 (83, 84) | Novel Protein sim. GBank gij3820584 (AF086791) - carbamoylphosphate synthetase large subunit [Zymomonas mobilis] | | synthase | 264602 | |
| 43 | 20467069 (85, 86) | Novel Protein sim. GBank gij3820584 (AF086791) - carbamoylphosphate synthetase large subunit [Zymomonas mobilis] | | | 264605 | |
| 44 | 13085287 (87, 88) | Novel Protein sim. GBank gij2494764[sp]Q50729[GUA_MYCTU - GMP SYNTHASE (GLUTAMINE-HYDROLYZING) (GLUTAMINE AMIDOTRANSFERASE) (GMP SYNTHETASE)] | Contains protein domain (PF00958) - GMP synthase C terminal domain | synthase | 264769, 264636 | |
| 45 | 39384711 (89, 90) | Novel Protein sim. GBank gij1881738 (U89688) - myosin-I binding protein Acan125 [Acanthamoeba castellanii] | | UNCLASSIFIED | 264769, 264510, 264508 | |
| 46 | 95003398 (91, 92) | | | ngl | 264566 | |
| 47 | 11698624 (93, 94) | | | UNCLASSIFIED | 264689 | |
| 48 | 79407218 (95, 96) | | | | 18108385, 264635, 264828 | |
| 49 | 21659844 (97, 98) | | | UNCLASSIFIED | 264603 | |
| 50 | 80503996 (99, 100) | | | | 264508, 264603, 264769, 264689, 264636, 264558, 264486 | |
| 51 | 80255569 (101, 102) | Novel Protein sim. GBank gij3411177 (AF076240) - MccC [Rhizobium leguminosarum bv. viciae] | | UNCLASSIFIED | 264593, 18108387 | |
| 52 | 78208528 (103, 104) | Novel Protein sim. GBank gij3914892[sp]Q26284[SM41_HEMPU - 41 KD SPICULE MATRIX PROTEIN PRECURSOR (HSM41) (HPSMC)] | | sinud | 264634 | |
| 53 | 36989670 (105, 106) | Novel Protein sim. GBank gij3880411 (AC004561) - putative proline-rich protein [Arabidopsis thaliana] | | UNCLASSIFIED | 264782 | |
| 54 | 79570897 (107, 108) | Novel Protein sim. GBank gij1633572 (U52064) - Herpesvirus saimiri ORF73 homolog [Kaposi's sarcoma-associated herpes-like virus] | | UNCLASSIFIED | 264630, 264909, 264766 | |
| 55 | 80202703 (109, 110) | Novel Protein sim. GBank gij4321580[gblAAD15785] - (AF050114) alginatase [Pseudomonas sp. W7] | | | 29331824, 264102, 265018, 18108376 | |
| 56 | 8758408 (111, 112) | | | | 264604 | |
| 57 | 11223386 (113, 114) | | Contains protein domain (PF00076) - RNA recognition motif (a.k.a. RRM, RBD, or RNP domain) | UNCLASSIFIED | 264557 | |

| | | | | | |
|----|---------------------|--|---|---------------|---|
| 58 | 91227508 (115, 116) | Novel Protein sim. GBank gij15616074 gb AAD45616.1 AF08194 - (AF081943) protease- derived STE20-like kinase PSK [Homo sapiens] | Contains protein domain (PF00069) - Eukaryotic protein kinase domain | kinase | 56182575, 264259, 60432049, 35896052, 66712502, 264909, 265008, 265010, 265011, 264681, 29148784, 35695917, 60170615, 264691, 264692, 264693, 18108374, 35696423, 56182323, 60432113 264600, 264689, 264638 |
| 59 | 80077371 (117, 118) | Novel Protein sim. GBank gij1172920 sp P45830 REFE_MYCLE - PUTATIVE UNDECAPRENYL-PHOSPHATE ALPHA-N- ACETYLGLUCOSAMINYLTRANSFERASE | Contains protein domain (PF00953) - Glycosyl transferase | transferase | |
| 60 | 12958341 (119, 120) | | | | 264689 |
| 61 | 80426808 (121, 122) | Novel Protein sim. GBank gij11710216 (U79260) - unknown [Homo sapiens] | | glycoprotein | 264766 |
| 62 | 13504866 (123, 124) | | | | 264630 |
| 63 | 16474553 (125, 126) | | | UNCLASSIFIED | 265019 |
| 64 | 20724578 (127, 128) | Novel Protein sim. GBank gij420945 pir JA47041 - transposase homolog (insertion element [SAE1]) - Alcaligenes eutrophus | | UNCLASSIFIED | 264602 |
| 65 | 79326308 (129, 130) | Novel Protein sim. GBank gij3123212 sp O06134 KPYK_MYCTU - PYRUVATE KINASE (PK) | Contains protein domain (PF00224) - Pyruvate kinase | kinase | 264563 |
| 66 | 46854384 (131, 132) | Novel Protein sim. GBank gij3928723 emb CAA22218 - (AL034355) putative ABC transporter (Streptomyces coelicolor) | | transport | 22278996, 264558 |
| 67 | 78952543 (133, 134) | Novel Protein sim. GBank gij231985 sp P30234 DHA_MYCTU - ALANINE DEHYDROGENASE (40 KD ANTIGEN) | | dehydrogenase | 265021 |
| 68 | 79817382 (135, 136) | | | | 264909 |
| 69 | 79841764 (137, 138) | | | UNCLASSIFIED | 264908 |
| 70 | 79871329 (139, 140) | | | | 264906, 264908 |
| 71 | 65897458 (141, 142) | | | UNCLASSIFIED | 264602, 265021 |
| 72 | 87734977 (143, 144) | Novel Protein sim. GBank gij4415926 gb AAD20157 - (AC006282) unknown protein [Arabidopsis thaliana] | | UNCLASSIFIED | 264488, 264905, 264906, 264907, 264908, 264511, 265008, 264910, 264758, 87168474, 264682, 264766, 264686, 264689, 35695917, 265021, 60170615, 264691, 33657023, 264692, 264693, 264629, 264631, 264639, 22279000 |
| 73 | 80025241 (145, 146) | | | UNCLASSIFIED | 60424179, 264508, 264908, 265007, 264603, 264687, 264689, 264692, 18108387 |
| 74 | 20377410 (147, 148) | | | UNCLASSIFIED | 264605 |
| 75 | 11819032 (149, 150) | Novel Protein sim. GBank gij2853098 emb CAA16914 - (AL021767) vacuolar protein sorting [Schizosaccharomyces pombe] | | UNCLASSIFIED | 264689 |
| 76 | 95105303 (151, 152) | Novel Protein sim. GBank gij446881 emb CAB38212 - (AL035601) putative protein [Arabidopsis thaliana] | | UNCLASSIFIED | 83373044, 264906, 264557 |
| 77 | 10144718 (153, 154) | Novel Protein sim. GBank gij854065 emb CAA58337 - (X83413) U88 [Human herpesvirus 6] | | UNCLASSIFIED | 264563 |
| 78 | 8758258 (155, 156) | | | UNCLASSIFIED | 264604 |

| | | | | | |
|----|---------------------|--|---|--|--|
| 79 | 94140190 (157, 159) | Novel Protein sim. GBank gij5689453[dj]BAA83010.1] - (AB028981) KIAA1058 protein [Homo sapiens] | Contains protein domain (PF00169) - PH domain | | 35696286, 22278998, 29331822, 29331824, 29331825, 29331827, 264905, 264906, 264907, 66712502, 264908, 264908, 265008, 265009, 264910, 60170831, 55812038, 33109954, 265017, 265018, 264288, 264768, 56181562, 21906765, 21906769, 29148784, 265020, 264690, 264691, 264692, 264693, 60431528, 35696423, 264631, 264632, 264634, 264636, 264639, 83373044, 264564, 264566, 264567 |
| 80 | 82314840 (159, 160) | | UNCLASSIFIED | | 264769, 264501, 265006, 264910, 264604, 264605, 264634, 264635, 264905, 264762, 264637, 264592, 264628, 264907, 264691, 264908, 264567, 264909, 264766 |
| 81 | 20467247 (161, 162) | Novel Protein sim. GBank gij1723442[sp]Q10258[YD2A, SCHPO - HYPOTHETICAL 69.0 KD PROTEIN C56F8.10 IN CHROMOSOME 1 | reductase | | 264605 |
| 82 | 16331388 (163, 164) | Novel Protein sim. GBank gij2895866 (AF045770) - methylmalonate semi-aldehyde dehydrogenase [Oryza sativa] | dehydrogenase | | 264567 |
| 83 | 94741180 (165, 166) | Novel Protein sim. GBank gij3402673 (AC004697) - unknown protein [Arabidopsis thaliana] | UNCLASSIFIED | | 264488, 264508, 264509, 264905, 264908, 264909, 264511, 264591, 264593, 264594, 264595, 264596, 264758, 264603, 264760, 264681, 18108351, 264762, 264682, 264764, 264684, 264766, 264686, 264632, 264637, 264557, 264638, 264639, 18108385, 264556 |
| 84 | 80355375 (167, 168) | Novel Protein sim. GBank gij1173364[sp]P45380[SAT1, RAT - SULFATE ANION TRANSPORTER 1 (CANALICULAR SULFATE TRANSPORTER) (SULFATE/CARBONATE ANTIporter) | transport | | 264508, 264908, 264907, 264908, 264909, 264910, 264760, 264763, 264764, 264766, 264768, 264769, 35695855, 264636, 264637 |
| 85 | 80498600 (169, 170) | Novel Protein sim. GBank gij2120998[sp]S70682 - glycosyltransferase homolog - Bordetella pertussis | transferase | | 264605, 264762, 264687, 264769, 18108374, 264636, 264486 |
| 86 | 39559043 (171, 172) | Novel Protein sim. GBank gij3256023[emb]CAA17228.1] - (AL021897) hypothetical protein Rv1112 [Mycobacterium tuberculosis] | | | 264910 |
| 87 | 13856808 (173, 174) | | UNCLASSIFIED | | 264093 |

| | | | | |
|----|---------------------|---|---|--|
| 88 | 95344718 (175, 176) | Novel Protein sim. GBank gjl559703[dbj BAA07552] - (D38549) ha1025 is new [Homo sapiens] | | 52644507, 52646365, 18108398, 65274572, 56182575, 56994075, 35696286, 22278997, 22278998, 22278999, 264092, 264093, 264094, 264095, 264259, 29331822, 29331824, 58182181, 66714117, 29331825, 29331826, 60432289, 29331827, 29331828, 35696052, 33656970, 264105, 264508, 264905, 264906, 264907, 264908, 29331830, 66712502, 52644045, 58182435, 285007, 265009, 264910, 60170831, 264592, 60431735, 60433358, 33657402, 264757, 60433438, 55812038, 264758, 21906754, 52646317, 33109954, 52644298, 87168474, 265011, 87168559, 264601, 265017, 265018, 264604, 265019, 264448, 264369, 264288, 264766, 52644229, 21906768, 21906767, 21906768, 21906769, 55811857, 35695917, 265020, 265021, 265022, 60170815, 52644150, 33657023, 65274620, 33657109, 27486261, 27486264, 33657349, 35695763, 264628, 263972, 18108374, 55810764, 35696423, 55811576, 65274791, 35695855, 60431850, 264636, 52644332, 58182323, 60170394, 83373044, 18108385, 18108387, 18108388, 56526486, 87168518, 60432113, 22279000, 22279002, 264482, 264564, 264486 |
| 89 | 80077389 (177, 178) | Novel Protein sim. GBank gjl1710383[sp P46352 RIPX_BACSU - PROBABLE INTEGRASE/RECOMBINASE RIPX | | 264600 |
| 90 | 82115999 (178, 180) | Novel Protein sim. GBank gjl249989[sp P76403 YEGQ_ECOLI - PUTATIVE PROTEASE IN BAER-OGRK INTERGENIC REGION | UNCLASSIFIED protease | 264760 265006 |
| 92 | 78554871 (183, 184) | Novel Protein sim. GBank gjl3367754[emb CAA20079] - (AL031155) hypothetical protein SC3A7.16c [Streptomyces coelicolor] | UNCLASSIFIED | 264691 |
| 93 | 80496778 (185, 186) | Novel Protein sim. GBank gjl2895095 (AF011337) - putative E1-E2 ATPase [Mus musculus] | ATPase_associated | 264907, 264908, 264910, 265009, 264605, 264769 |
| 94 | 79646649 (187, 188) | Novel Protein sim. GBank gjl171919[sp P46920 OPUA_BACSU - GLYCINE BETAINES TRANSPORT ATP-BINDING PROTEIN OPUA | Contains protein domain (PF00571) - transport CBS domain | 284906 |
| 95 | 11090238 (189, 190) | | | 264594 |

| | | | | | |
|-----|---------------------|--|--|-------------------|--|
| 98 | 94322125 (191, 192) | Novel Protein sim. GBank gi 4589560 dbj BAA76802.1 - (AB023175) KIAA0958 protein [Homo sapiens] | | UNCLASSIFIED | 22278995, 22278999, 264259, 29331822, 29331826, 35696052, 28146499, 284509, 284906, 284907, 264908, 284909, 265007, 265008, 284910, 265009, 264593, 265010, 265017, 264604, 265019, 18108351, 264288, 264768, 264768, 264769, 21906765, 21906767, 21905769, 265020, 264692, 33657182, 35695763, 264628, 264629, 18108379, 284631, 284636, 18108381, 264559, 18108382, 83373044, 22278002, 284508 |
| 97 | 79605200 (193, 194) | Novel Protein sim. GBank gi 4583559 emb CAB40388.1 - (AJ005255) OxyR [Erwinia chrysanthemi] | | UNCLASSIFIED | |
| 96 | 79427000 (195, 196) | Novel Protein sim. GBank gi 1001693 dbj BAA10430 - (D64002) hypothetical protein [Synechocystis sp.] | | UNCLASSIFIED | 264909 |
| 99 | 20466324 (197, 198) | Novel Protein sim. GBank gi 1169479 sp P43925 EFG_HAEIN - ELONGATION FACTOR G (EF-G) | | UNCLASSIFIED | 264605 |
| 100 | 79840113 (199, 200) | Novel Protein sim. GBank gi 480897 pir J37485 - gene | | UNCLASSIFIED | 264693 |
| 101 | 80203298 (201, 202) | msg1 protein - mouse | | UNCLASSIFIED | 265020, 264102, 263972 |
| 102 | 20467259 (203, 204) | Novel Protein sim. GBank gi 2894166 emb CAA11773.1 - (AJ223998) PCZA361.18 [Amycolatopsis orientalis] | | synthase | 264605 |
| 103 | 20466368 (205, 206) | Novel Protein sim. GBank gi 1731040 sp P54509 YQH_H_BACSU - HYPOTHETICAL HELICASE IN SINI-GCVT INTERGENIC REGION | Contains protein domain (PF00271) - helicase | | 264605 |
| 104 | 80247572 (207, 208) | Novel Protein sim. GBank gi 854055 emb CAA58337 - (X83413) U88 [Human herpesvirus 6] | | UNCLASSIFIED | 264591, 264595, 264602 |
| 105 | 79605206 (209, 210) | Novel Protein sim. GBank gi 1685117 U70770 - furrowed [Drosophila melanogaster] | Contains protein domain (PF00084) - complement | | 264508 |
| 106 | 28382058 (211, 212) | Novel Protein sim. GBank gi 1705505 sp P54729 BS4_MOUSE - BS4 PROTEIN | Sushi domain (SCR repeat) | UNCLASSIFIED | 264511, 265009 |
| 107 | 80057791 (213, 214) | Novel Protein sim. GBank gi 4887229 gb AAD32244.1 AF150755 - (AF150755) microtubule-actin crosslinking factor [Mus musculus] | UBA domain | ATPase-associated | 29331824, 264591, 21906754, 265019 |
| 108 | 80237636 (215, 216) | Novel Protein sim. GBank gi 263577 emb CAB15264 - (Z99120) similar to ABC transporter (ATP-binding protein) [Bacillus subtilis] | Contains protein domain (PF00005) - transport | | 18108374, 35695917, 22278996, 264113, 264600, 264602, 264603, 265017, 264910, 264906, 264636, 264766 |
| 109 | 95194148 (217, 218) | Novel Protein sim. GBank gi 233079 emb CAB11265 - (Z98601) carboxypeptidase s precursor [Schizosaccharomyces pombe] | ABC transporter | UNCLASSIFIED | 264758, 264603, 264630, 264636, 264637 |
| 110 | 79582823 (219, 220) | | | | 264687 |
| 111 | 39565458 (221, 222) | | | | 264564 |
| 112 | 79856038 (223, 224) | | | UNCLASSIFIED | 264908 |
| 113 | 17959439 (225, 226) | | | UNCLASSIFIED | 265007 |
| 114 | 80502101 (227, 228) | | | UNCLASSIFIED | 264769 |

| | | | | |
|-----|---------------------|--|---|--|
| 115 | 80251003 (229, 230) | Novel Protein sim. GBank gij2246532 (U83872) - ORF 73, contains large complex repeat CR 73 [Kaposi's sarcoma-associated herpesvirus] | UNCLASSIFIED | 52645156, 52645080, 33656970, 264592, 21908754, 27486264, 18108379, 35696423, 264635, 52644332, 18108382 |
| 116 | 81286689 (231, 232) | | | 264905, 264906, 264907, 264908, 264909, 264910, 264758, 265010, 264763, 264682, 264764, 264766, 264885, 264886, 264768, 264769, 33657023, 264693, 33657109, 264628, 18108374, 264631, 264632, 264634, 264636, 264637, 264638, 264639, 56528486, 264585, 264586 |
| 117 | 78636695 (233, 234) | | UNCLASSIFIED | 264639, 264693 |
| 118 | 80222170 (235, 236) | | Contains protein domain (PF00170) - bZIP transcription factor | 263974 |
| 119 | 91013071 (237, 238) | Novel Protein sim. GBank gij732526 (U22327) - alpha2(IV) collagen [Caenorhabditis elegans] | UNCLASSIFIED | 22278996, 29331824, 60432289, 265007, 60433438, 264603, 264605, 18108351, 264769, 264689, 265020, 264534, 27486261, 264558, 83373044, 18108385, 264564 |
| 120 | 8756481 (239, 240) | Novel Protein sim. GBank gij2131219 [pir] [S50157] - cyclin-dependent kinase chain SRB10 - yeast (Saccharomyces cerevisiae) | kinase | 264603 |
| 121 | 80026153 (241, 242) | | | 264595 |
| 122 | 20457620 (243, 244) | Novel Protein sim. GBank gij2052147 [emb] [CAB08137] - (Z94752) ksgA [Mycobacterium tuberculosis] | transferase | 264605 |
| 123 | 8758278 (245, 246) | | Ribosomal RNA adenine dimethylases | |
| 124 | 79104017 (247, 248) | Novel Protein sim. GBank gij2833385 [sp] [Q43134] [UGST_SORBI - GRANULE-BOUND GLYCOPROTEIN (STARCH) SYNTHASE PRECURSOR | synthase | 264604 |
| 125 | 87797988 (249, 250) | Novel Protein sim. GBank gij475542 (U08255) - glutamate receptor delta-1 subunit [Rattus norvegicus] | | 18108394, 18108397, 265006, 265007, 265008, 265010, 265011, 18108355, 18108379, 18108380, 18108384 |
| 126 | 56701283 (251, 252) | Novel Protein sim. GBank gij5102785 [emb] [CAB45200.1] - (AL078308) putative transcriptional regulator [Streptomyces coelicolor] | misc_channel | 264508, 264906, 265009, 264596, 22278002 |
| 127 | 20467267 (253, 254) | | | 264511 |
| 128 | 80248473 (255, 256) | Novel Protein sim. GBank gij130120 [sp] [P23620] [PHOB_PSEAE - PHOSPHATE REGULON TRANSCRIPTIONAL REGULATORY PROTEIN PHOB | UNCLASSIFIED | 264605 |
| 129 | 95280543 (257, 258) | Novel Protein sim. GBank gij2508493 [sp] [P38036] [YGC8_ECOLI - HYPOTHETICAL 100.5 KD PROTEIN IN IAP-CYSH INTERGENIC REGION | phosphatase | 264907, 264909, 264910, 264600, 264601, 264603, 264605, 18108351, 264693, 264557 |
| 130 | 80085583 (259, 260) | Novel Protein sim. GBank gij854065 [emb] [CAA58337] - (X83413) U88 [Human herpesvirus 6] | UNCLASSIFIED | 35696423, 35695855, 264600, 264602, 264603, 264604, 264605, 264508, 264908, 264564, 264628, 264682, 264555, 264683 |
| 131 | 94895022 (261, 262) | Novel Protein sim. GBank gij1076038 [pir] [S54860 - ABC transporter PstC-2 chain - Mycobacterium tuberculosis] | transport | 264634 |

| | | | | | |
|-----|---------------------|---|---|---------------|--|
| 132 | 10887692 (263, 264) | Novel Protein sim. GBank gij1877340[emb]CAB07068] - (Z92771) accA3 [Mycobacterium tuberculosis] | Contains protein domain (PF00289) Carbamoyl-phosphate synthase (CPSase) | carboxylase | 264636 |
| 133 | 94630883 (265, 266) | Novel Protein sim. GBank gij4585838[emb]CAB40932.1] - (AL049630) putative NADH dehydrogenase [Streptomyces coelicolor] | | dehydrogenase | 264905, 264605, 265021 |
| 134 | 79834860 (267, 268) | Novel Protein sim. GBank gij1460074[emb]CAB01049] - (Z77250) hypothetical protein Rv2566 [Mycobacterium tuberculosis] | | | 264634 |
| 135 | 19885057 (269, 270) | Novel Protein sim. GBank gij2125896[emb]CAA73511] - (Y13070) isopolyglutamate synthase [Streptomyces coelicolor] | | synthase | 264508 |
| 136 | 79846083 (271, 272) | | | | |
| 137 | 79819770 (273, 274) | | | | 264683, 264685, 264686, 264691, 264692, 264693 |
| 138 | 79635971 (275, 276) | Novel Protein sim. GBank gij5420387[emb]CAB46679.1] - (AJ243459) proteophosphoglycan [Leishmania major] | | UNCLASSIFIED | 18108374, 18108385, 33657109, 33657182, 265010, 22278998, 265006, 265007, 265008, 265009, 264693 |
| 139 | 86688076 (277, 278) | Novel Protein sim. GBank gij5689912[emb]CAB52075.1] - (AL109732) putative mutase [Streptomyces coelicolor A3(2)] | Contains protein domain (PF01817) Chorismate mutase | dehydrogenase | 22278996, 265007, 264910, 60433356, 265010, 264602, 264605, 264766, 264688, 264769, 264693, 32833986, 18108374, 18108387 |
| 140 | 79825759 (279, 280) | | | UNCLASSIFIED | 264908 |
| 141 | 20700094 (281, 282) | | | | 264600 |
| 142 | 80028104 (283, 284) | Novel Protein sim. GBank gij3581916[emb]CAA20855] - (AL031545) mutS family DNA mismatch repair protein [Schizosaccharomyces pombe] | | nuclease | 264602, 265017 |
| 143 | 11072274 (285, 286) | | | | |
| 144 | 95009102 (287, 288) | Novel Protein sim. GBank gij3334127[sp]P97303[BAC2_MOUSE - TRANSCRIPTION REGULATOR PROTEIN BACH2 (BTB AND CNC HOMOLOG 2)] | | UNCLASSIFIED | 264600 |
| 145 | 80027058 (289, 290) | Novel Protein sim. GBank gij3757569[emb]CAA21315] - (AL031863) 1-evidence=predicted by content; 1-method=genefinder;084; 1-method_score=66.31; 1-evidence_end [Drosophila melanogaster] | | | 263978, 264600, 264910, 264632, 264508, 264563, 264564, 264591, 264556, 264908, 264629, 264639 |
| 146 | 13085662 (291, 292) | Novel Protein sim. GBank gij140807[sp]P24536[Y121_BURCE - INSERTION ELEMENT IS402 HYPOTHETICAL 24 KD PROTEIN] | Contains protein domain (PF01675) Transposase | | 22278996, 264602 |
| 147 | 94320366 (293, 294) | Novel Protein sim. GBank gij2827608[emb]CAA16663] - (AL021646) uvrD2 [Mycobacterium tuberculosis] | | helicase | 264905, 264906, 264809, 264510, 265009, 60433356, 264600, 264601, 264804, 264805, 264687, 264769, 18108365, 65274791, 18108387 |
| 148 | 80248804 (295, 296) | Novel Protein sim. GBank gij2916947[emb]CAA17585] - (AL021999) hypothetical protein Rv0986 [Mycobacterium tuberculosis] | | transport | 265009, 265010, 264600, 264602, 264603, 264604, 264605, 264693, 33657109, 264636 |

| | | | | |
|-----|---------------------|---|---|---|
| 149 | 80249373 (297, 298) | Novel Protein sim. GBank gij1723073 sp Q11040Y081_MYCTU - HYPOTHETICAL ABC TRANSPORTER ATP-BINDING PROTEIN CY50.01 | Contains protein domain (PF00005) - transport ABC transporter | 265010, 264600, 264601, 264603, 264604, 27486285, 264636 |
| 150 | 20284748 (289, 300) | Novel Protein sim. GBank gij3724125 emb CAA11905 - (AJ224340) maltose phosphorylase [Lactobacillus sanfrancisco] | | 264600 |
| 151 | 20726398 (301, 302) | Novel Protein sim. GBank gij729312 sp P07651 DEOB_ECOLI - PHOSPHOENOMUTASE (PHOSPHODEOXYRIBOMUTASE) | Contains protein domain (PF01676) - Metalloenzyme superfamily | 264602 |
| 152 | 95002877 (303, 304) | Novel Protein sim. GBank gij2497952 sp P55667 Y4TM_RHISN - HYPOTHETICAL HYDROLASE/PEPTIDASE Y4TM | peptidase | 264602 |
| 153 | 80256665 (305, 306) | Novel Protein sim. GBank gij3123021 sp Q90508 VIT1_FUNHE - VITELLOGENIN I PRECURSOR (VTG I) (CONTAINS: LIPOVITELLIN 1 (LV1); PHOSVITIN (PV); LIPOVITELLIN 2 (LV2)) | UNCLASSIFIED | 264593 |
| 154 | 82305966 (307, 308) | Novel Protein sim. GBank gij418697 pir JN0443 - transcription initiation factor sigma homolog hdb - Streptomyces aureofaciens | Contains protein domain (PF00140) - Sigma-70 factor | 264910, 264762, 264691, 264634 264605 |
| 155 | 20428859 (309, 310) | Novel Protein sim. GBank gij628710 pir IS41739 - hypothetical protein - Escherichia coli | UNCLASSIFIED | 264585 |
| 156 | 39564742 (311, 312) | Novel Protein sim. GBank gij3695013 (AF052586) - CIRA (Pseudomonas aeruginosa) | Contains protein domain (PF00142) - 4Fe-4S iron sulfur cluster binding proteins, NifH/HnfC family | 264691 |
| 157 | 10358887 (313, 314) | Novel Protein sim. GBank gij1073072 pir C55543 - cmaU protein - Pseudomonas syringae pv. syringae | UNCLASSIFIED | 264805 |
| 158 | 79761936 (315, 316) | Novel Protein sim. GBank gij1173023 sp P46789 RL30_STRCO - 50S RIBOSOMAL PROTEIN L30 | Contains protein domain (PF00400) - WD domain, G-beta repeat | 265008 264605 |
| 159 | 78890376 (317, 318) | Novel Protein sim. GBank gij5304869 emb CAB46028.1 - (AL031685) dJ963K23.2 (novel protein) [Homo sapiens] | Contains protein domain (PF00327) - Ribosomal protein L30p/L7e | 22278996, 264600, 264603, 35695917, 32833986, 35696423, 264636 |
| 160 | 11075119 (319, 320) | Novel Protein sim. GBank gij76177 pir Q9ECFT - hypothetical 38.8K protein (ftsI 5' region) - Escherichia coli | interleukin Zinc finger, C3HC4 type (RING finger) | 264112, 264532, 22278002 |
| 161 | 80055007 (321, 322) | Novel Protein sim. GBank gij1173023 sp P46789 RL30_STRCO - 50S RIBOSOMAL PROTEIN L30 | UNCLASSIFIED | 264639 |
| 162 | 80016371 (323, 324) | Novel Protein sim. GBank gij5304869 emb CAB46028.1 - (AL031685) dJ963K23.2 (novel protein) [Homo sapiens] | UNCLASSIFIED | 264905, 264907, 264600 |
| 163 | 11692308 (325, 326) | Novel Protein sim. GBank gij2661691 emb CAA15795 - (AL009204) putative protease [Streptomyces coelicolor] | UNCLASSIFIED | 264691 |
| 164 | 80077902 (327, 328) | Novel Protein sim. GBank gij4416478 gb AAD20378 - (AF125999) transposase [Mycobacterium avium] | UNCLASSIFIED | 264605, 264486 |
| 165 | 10856067 (329, 330) | Novel Protein sim. GBank gij76177 pir Q9ECFT - hypothetical 38.8K protein (ftsI 5' region) - Escherichia coli | UNCLASSIFIED | 265010 |
| 166 | 88095003 (331, 332) | Novel Protein sim. GBank gij76177 pir Q9ECFT - hypothetical 38.8K protein (ftsI 5' region) - Escherichia coli | UNCLASSIFIED | 264600 |
| 167 | 16395460 (333, 334) | Novel Protein sim. GBank gij76177 pir Q9ECFT - hypothetical 38.8K protein (ftsI 5' region) - Escherichia coli | UNCLASSIFIED | 264556, 264557, 264558, 264559 |
| 168 | 80079362 (335, 336) | Novel Protein sim. GBank gij76177 pir Q9ECFT - hypothetical 38.8K protein (ftsI 5' region) - Escherichia coli | UNCLASSIFIED | |
| 169 | 80239581 (337, 338) | Novel Protein sim. GBank gij76177 pir Q9ECFT - hypothetical 38.8K protein (ftsI 5' region) - Escherichia coli | UNCLASSIFIED | |

| | | | | | |
|-----|---------------------|---|--|--|---------------|
| 170 | 78612364 (339, 340) | Novel Protein sim. GBank gij140888[sp]27847[YIGK_ECOLI - HYPOTHETICAL 15.4 KD PROTEIN IN RECQ-PLDB INTERGENIC REGION (F138)] | Contains protein domain (PF01810) - LysE type translocator | 264906 | |
| 171 | 95293073 (341, 342) | | | 264595, 264604 | |
| 172 | 37797007 (343, 344) | Novel Protein sim. GBank gij4210905[gb]AAD12048.1] - (AF045609) AglG [Sinorhizobium meliloti] | Contains protein domain (PF00528) - Binding-protein-dependent transport systems inner membrane component | 264769 | transport |
| 173 | 57529660 (345, 346) | Novel Protein sim. GBank gij132854[sp]P02387[RL2_ECOLI - 50S RIBOSOMAL PROTEIN L2] | Contains protein domain (PF00181) - Ribosomal Proteins L2 | 264769 | ribosomalprot |
| 174 | 95293078 (347, 348) | Novel Protein sim. GBank gij1881350[dbj]BAA193771 - (AB001488) PROBABLE TRANSPORT PROTEIN, SIMILAR TO ANTIOTIC TRANSPORT-ASSOCIATED PROTEIN ACTII IN STREPTOMYCES COELICOLOR. [Bacillus subtilis] | transport | 264510, 264593, 264602, 264603, 264605, 264762, 264693 | |
| 175 | 79758270 (349, 350) | Novel Protein sim. GBank gij207272[emb]CAB08326] - (Z95121) manA [Mycobacterium tuberculosis] | isomerase | 264565 | |
| 176 | 80066896 (351, 352) | Novel Protein sim. GBank gij1055198 (U40187) - similar to PIR:A41724 chicken LD (limb deformity) gene product and to formin; also P-rich region similar to collagen [Caenorhabditis elegans] | UNCLASSIFIED | 264907, 264910, 264681, 264558 | |
| 177 | 86684852 (353, 354) | Novel Protein sim. GBank gij232873[emb]CAB10952] - (Z98268) hypothetical protein Rv1695 [Mycobacterium tuberculosis] | Contains protein domain (PF01513) - Domain of unknown function | 264768, 60424179, 264687, 264688, 264769, 29331826, 60432289, 18108376, 264689, 18108387, 32833986, 22278996, 265020, 264600, 264601, 264602, 264603, 264604, 264805, 264635, 264762, 264636, 264908, 264564, 264637, 264638, 264486, 60433356, 264766 | |
| 178 | 78559526 (355, 356) | Novel Protein sim. GBank gij1906596 (U81788) - kinesin-73 [Drosophila melanogaster] | struct | 264693, 33657109, 264635 | |
| 179 | 20263112 (357, 358) | | UNCLASSIFIED | 264563 | |
| 180 | 80488958 (359, 360) | Novel Protein sim. GBank gij1169387[sp]P45256[DNAB_HAEIN - REPLICATIVE DNA HELICASE] | helicase | 264769 | |
| 181 | 78585369 (361, 362) | Novel Protein sim. GBank gij3170615 (AF059485) - DOC4 [Mus musculus] | UNCLASSIFIED | 21906767, 264635, 264639, 18108384 | |
| 182 | 80577899 (363, 364) | | UNCLASSIFIED | 264259, 35696052, 56182435, 264511, 265018, 33657109, 264555, 264568 | |
| 183 | 11614017 (365, 366) | Novel Protein sim. GBank gij1076627[pir]S54172 - inorganic pyrophosphatase (EC 3.6.1.1) - common tobacco | UNCLASSIFIED | 264690 | |
| 184 | 10174167 (367, 368) | Novel Protein sim. GBank gij4371280[gb]AAD18138] - (AC006260) hypothetical protein [Arabidopsis thaliana] | UNCLASSIFIED | 264510 | |

| | | | | | |
|-----|---------------------|--|---|--------------|--|
| 185 | 21660822 (369, 370) | Novel Protein sim. GBank gij3006178[emb]CAA18398.1] - (AL02304) putative mma transport regulator [Schizosaccharomyces pombe] | | UNCLASSIFIED | 264604 |
| 186 | 80070329 (371, 372) | Novel Protein sim. GBank gij289802[sp]P94408[YCLF_BACSU - HYPOTHETICAL 53.3 KD PROTEIN IN SFP-GERKA INTERGENIC REGION] | | transport | 264595 |
| 187 | 80186611 (373, 374) | Novel Protein sim. GBank gij3150260[emb]CAA19179] - (AL023634) cyclin [Schizosaccharomyces pombe] | | UNCLASSIFIED | 264369 |
| 188 | 20464942 (375, 376) | Novel Protein sim. GBank gij2145853[pi]S72938 - hix protein - Mycobacterium leprae | | kinase | 264605 |
| 189 | 82338215 (377, 378) | Novel Protein sim. GBank gij1881244[db]BAA19271] - (AB001488) SIMILAR TO PYRUVATE OXIDASE AND ACETOLACTATE SYNTHASE. [Bacillus subtilis] | | UNCLASSIFIED | 35896052, 264602, 264605, 264762, 264689, 35895917, 18108370, 18108372, 264636, 264565 |
| 190 | 80086821 (379, 380) | Novel Protein sim. GBank gij120226[sp]P28725[FKBP_STRCH - FK506-BINDING PROTEIN (PEPTIDYL-PROLYL CIS-TRANS ISOMERASE) (PIPIASE) (ROTAMASE)] | Contains protein domain (PF00205) - Thiamine pyrophosphate enzymes | synthase | 284563 |
| 191 | 88095012 (381, 382) | Novel Protein sim. GBank gij1705461[sp]P53656[BIOA_ERWHE - ADENOSYLMETHIONINE-8-AMINO-7-OXONONANOATE AMINOTRANSFERASE (7.8-DIAMINO-PELAGONIC ACID AMINOTRANSFERASE) (DAPA AMINOTRANSFERASE)] | Contains protein domain (PF00254) - FKBP-type peptidyl-prolyl cis-trans isomerases | isomerase | 264508, 264604, 264605, 264769, 264555 |
| 192 | 16333378 (383, 384) | Novel Protein sim. GBank gij4980892[gb]AAD35474.1(AE00171 - (AE001718) ABC transporter, ATP-binding protein [Thermotoga maritima] | | | 264567 |
| 193 | 78910127 (385, 386) | Novel Protein sim. GBank gij1705461[sp]P53656[BIOA_ERWHE - ADENOSYLMETHIONINE-8-AMINO-7-OXONONANOATE AMINOTRANSFERASE (7.8-DIAMINO-PELAGONIC ACID AMINOTRANSFERASE) (DAPA AMINOTRANSFERASE)] | | | 264908, 264693 |
| 194 | 20464949 (387, 388) | Novel Protein sim. GBank gij4980892[gb]AAD35474.1(AE00171 - (AE001718) ABC transporter, ATP-binding protein [Thermotoga maritima] | | | 264605 |
| 195 | 13518389 (389, 390) | Novel Protein sim. GBank gij1705461[sp]P53656[BIOA_ERWHE - ADENOSYLMETHIONINE-8-AMINO-7-OXONONANOATE AMINOTRANSFERASE (7.8-DIAMINO-PELAGONIC ACID AMINOTRANSFERASE) (DAPA AMINOTRANSFERASE)] | | transport | 264636 |
| 196 | 95005569 (391, 392) | Novel Protein sim. GBank gij3122305[sp]Q27778[K6PF_SCHMA - 6- PHOSPHOFRUCTOKINASE (PHOSPHOFRUCTOKINASE) (PHOSPHOHEXOKINASE)] | Contains protein domain (PF00202) - Aminotransferases class-III pyridoxal phosphate | gaba | 264600, 264689, 264638 |
| 197 | 80248665 (393, 394) | Novel Protein sim. GBank gij3122305[sp]Q27778[K6PF_SCHMA - 6- PHOSPHOFRUCTOKINASE (PHOSPHOFRUCTOKINASE) (PHOSPHOHEXOKINASE)] | Contains protein domain (PF00365) - Kinase | kinase | 264602, 264882, 264692, 18108374 |
| 198 | 79163635 (395, 396) | Novel Protein sim. GBank gij1781203[emb]CAB06110] - (Z83859) gnd [Mycobacterium tuberculosis] | | | 264636 |
| 199 | 78890715 (397, 398) | Novel Protein sim. GBank gij2642222 (AF030885) - telomere-associated recQ-like helicase [Usilago maydis] | Contains protein domain (PF00393) - 6-phosphogluconate dehydrogenases | | 265008 |
| 200 | 78413849 (399, 400) | Novel Protein sim. GBank gij2642222 (AF030885) - telomere-associated recQ-like helicase [Usilago maydis] | | UNCLASSIFIED | 264595, 264596 |
| 201 | 86945924 (401, 402) | Novel Protein sim. GBank gij2894379[emb]CAA74911.1] - (Y14573) ring finger protein [Hordeum vulgare] | | UNCLASSIFIED | 29331826, 265007, 264512, 33657402, 264596, 265017, 18108351, 264682, 264683, 264767, 264629, 55810764, 264634, 264635, 56182323, 60432113, 22279000 |

| | | | | | |
|-----|---------------------|---|--|--|--|
| 202 | 79588046 (403, 404) | Novel Protein sim. GBank gi231772 sp P30598 CHS1_USTMA - CHITIN SYNTHASE 1 (CHITIN-UDP ACETYL-GLUCOSAMINYL TRANSFERASE 1) | Contains protein domain (PF01644) - synthase Chitin synthase | 264600 | |
| 203 | 79843927 (405, 406) | Novel Protein sim. GBank gi1504042 dbj BAA13220 - (D86984) similar to yeast adenylate cyclase (S56776) [Homo sapiens] | | 22278995, 29331822, 29331825, 29331827, 264906, 21908754, 264683, 21908766, 21908769, 35696423, 264558 | |
| 204 | 79855186 (407, 408) | | UNCLASSIFIED | 264909 | |
| 205 | 10090563 (409, 410) | Novel Protein sim. GBank gi2633808 emb CAB13310 - (Z89111) similar to hypothetical proteins [Bacillus subtilis] | transport | 264909 | |
| 206 | 8758473 (411, 412) | | UNCLASSIFIED | 264604 | |
| 207 | 20754522 (413, 414) | Novel Protein sim. GBank gi2134381 pir IS0678 - polybromo 1 protein - chicken | UNCLASSIFIED | 264556 | |
| 208 | 20289261 (415, 416) | | | 264605 | |
| 209 | 80071069 (417, 418) | Novel Protein sim. GBank gi2501040 sp O05814 SYP_MYCTU - PROLYL-TRNA SYNTHETASE (PROLINE--TRNA LIGASE) (PRORS) | | 264605, 264689 | |
| 210 | 80168800 (419, 420) | | | 264905, 264907, 264909, 264766, 264687, 264691, 264629, 18108374, 264638 | |
| 211 | 80034539 (421, 422) | | | 263978 | |
| 212 | 82442474 (423, 424) | Novel Protein sim. GBank gi5031809 ref NP_005536.1 pISLR - immunoglobulin superfamily containing leucine-rich repeat | UNCLASSIFIED | 264508, 264905, 264906, 264907, 264908, 264600, 264762, 264534, 264632, 264634, 264635, 264639, 264486 | |
| 213 | 80248562 (425, 426) | Novel Protein sim. GBank gi3122359 sp O33123 LEU2_MYCLE - 3- ISOPROPYL MALATE DEHYDRATASE LARGE SUBUNIT (ISOPROPYL MALATE ISOMERASE) (ALPHA-IPM ISOMERASE) (IPMI) | isomerase | 22278996, 264508, 264602, 264603, 264605, 33657023, 264565, 264486 | |
| 214 | 80079381 (427, 428) | Novel Protein sim. GBank gi116236 sp P19421 CH60_COXBU - 60 KD CHAPERONIN (PROTEIN CPN60) (GROEL PROTEIN) (HEAT SHOCK PROTEIN B) | Contains protein domain (PF00118) - eph TCP-1/cpn60 chaperonin family | 264600, 264693 | |
| 215 | 14973283 (429, 430) | | | | |
| 216 | 80177716 (431, 432) | Novel Protein sim. GBank gi2506924 sp P49754 VP41_HUMAN - VACUOLAR ASSEMBLY PROTEIN VPS41 HOMOLOG (S53) | UNCLASSIFIED | 264629 | |
| 217 | 79603634 (433, 434) | Novel Protein sim. GBank gi1173288 sp P38108 IRSEA_ECOLI - SIGMA-E FACTOR NEGATIVE REGULATORY PROTEIN | dna_ma_bind | 264448 | |
| 218 | 80258475 (435, 436) | Novel Protein sim. GBank gi1173288 sp P38108 IRSEA_ECOLI - SIGMA-E FACTOR NEGATIVE REGULATORY PROTEIN | mapolymerase | 264594 | |
| 219 | 20438797 (437, 438) | Novel Protein sim. GBank gi1781097 emb CAB06231 - (Z83864) giB [Mycobacterium tuberculosis] | synthase | 264604 | |
| 220 | 13499572 (439, 440) | Novel Protein sim. GBank gi2964703 (AF052427) - unknown [Trypanosoma cruzi] | nucleaseinhib | 264689 | |
| 221 | 11287498 (441, 442) | Novel Protein sim. GBank gi4587313 dbj BAA76708.1 - (AB025248) alpha-1,2-mannosidase [Bacillus sp. M-90] | UNCLASSIFIED | 264555 | |

| | | | | | |
|-----|---------------------|--|---|--------------|--|
| 222 | 78862802 (443, 444) | Novel Protein sim. GBank gij1877269[emb]CAB07049] - (Z82770) hypothetical protein Rv0143c [Mycobacterium tuberculosis] | | UNCLASSIFIED | 264605, 264769, 35696423 |
| 223 | 83053869 (445, 446) | | | UNCLASSIFIED | 264908, 264907, 264603 |
| 224 | 79557920 (447, 448) | | | UNCLASSIFIED | 264684, 264693 |
| 225 | 79559541 (448, 450) | | | UNCLASSIFIED | 264692 |
| 226 | 79172397 (451, 452) | Novel Protein sim. GBank gij2274851[dbj]BAA21515] - (D64159) 3-7 gene product [Homo sapiens] | | UNCLASSIFIED | 22278998, 264112, 33657023, 263981 |
| 227 | 81777196 (453, 454) | Novel Protein sim. GBank gij868245 (U29488) - C56C10.7 gene product [Caenorhabditis elegans] | | UNCLASSIFIED | 35695917, 264636, 264907 |
| 228 | 79872285 (455, 456) | | | UNCLASSIFIED | 264768, 264907, 264908, 264692, 264593, 264639 |
| 229 | 79838266 (457, 458) | | | UNCLASSIFIED | 264908, 264910 |
| 230 | 11013209 (459, 460) | | | UNCLASSIFIED | 264631 |
| 231 | 20822207 (461, 462) | Novel Protein sim. GBank gij1835114[emb]CAA71733] - (Y10744) homoserine O-acetyltransferase [Leptospira meyeri] | | UNCLASSIFIED | 264906, 264600, 264603, 264692 |
| 232 | 80055035 (463, 464) | | | UNCLASSIFIED | 264600, 264603, 264605, 264687, 264769 |
| 233 | 80063054 (465, 466) | Novel Protein sim. GBank gij2642340 (AF032970) - imidazolone propionate hydrolase [Pseudomonas putida] | Contains protein domain (PF00449) - Urease | UNCLASSIFIED | 264604 |
| 234 | 7523998 (467, 468) | Novel Protein sim. GBank gij3510505 (AF030881) - polypeptide [Fugu rubripes] | | UNCLASSIFIED | 264369 |
| 235 | 80203671 (469, 470) | | | UNCLASSIFIED | 264106 |
| 236 | 78940001 (471, 472) | Novel Protein sim. GBank gij2104609[emb]CAB08605] - (Z95398) PckA [Mycobacterium leprae] | | UNCLASSIFIED | 264905 |
| 237 | 11755273 (473, 474) | | | UNCLASSIFIED | 264681 |
| 238 | 79461401 (475, 476) | | | UNCLASSIFIED | 264639 |
| 239 | 82435190 (477, 478) | Novel Protein sim. GBank gij2495617[sp]Q57252[YDIJ_HAEIN - HYPOTHETICAL PROTEIN H11163] | Contains protein domain (PF00037) - 4Fe-4S ferredoxins and related iron-sulfur cluster binding domains. | UNCLASSIFIED | 264906, 265010, 264603, 264762, 264682, 264636, 264638, 264486 |
| 240 | 21635375 (479, 480) | Novel Protein sim. GBank gij3183458[sp]P75796[YLIA_ECOLI - HYPOTHETICAL ABC TRANSPORTER ATP-BINDING PROTEIN YLIA] | transport | UNCLASSIFIED | 264259, 264769 |
| 241 | 80377307 (481, 482) | Novel Protein sim. GBank gij3875920[emb]CAB0411] - (Z81503) predicted using GeneFinder; similar to collagen; cDNA EST EMBL:D65450 comes from this gene; cDNA EST EMBL:D68888 comes from this gene [Caenorhabditis elegans] | | UNCLASSIFIED | 264908, 264909, 264764, 264639 |
| 242 | 82148454 (483, 484) | | | UNCLASSIFIED | 264489, 264907, 264908, 264511, 264760, 264764, 264692, 264635, 264637 |
| 243 | 79633207 (485, 486) | | | UNCLASSIFIED | 264906 |
| 244 | 80248582 (487, 488) | Novel Protein sim. GBank gij2624302[emb]CAA15575] - (AL008967) ald [Mycobacterium tuberculosis] | | UNCLASSIFIED | 264600, 264602, 264605, 264769, 264689 |
| 245 | 79863543 (489, 490) | Novel Protein sim. GBank gij2920825 (AF044499) - vgrE protein [Escherichia coli] | | UNCLASSIFIED | 264907, 264758 |
| 246 | 79162929 (491, 492) | Novel Protein sim. GBank gij5420387[emb]CAB46679.1] - (AJ243459) proteophosphoglycan [Leishmania major] | Contains protein domain (PF01106) - NiU-like domain | | 264637, 18108381, 18108387, 264565 |

| | | | | | |
|-----|---------------------|---|--|--------------|--|
| 247 | 79873185 (493, 494) | Novel Protein sim. GBank gij1838006[embj CA806648] - (Z85982) argB [Mycobacterium tuberculosis] | | kinase | 264909, 264691, 35696423, 18108387 |
| 248 | 80488883 (495, 496) | Novel Protein sim. GBank gij1168574[sp P42464 ATPB_CORGL - ATP SYNTHASE BETA CHAIN | | synthase | 35698286, 264907, 264511, 264602, 264768, 264688, 265021, 35695855, 18108385 |
| 249 | 78764845 (497, 498) | | | UNCLASSIFIED | 264907, 264910, 265011, 264762, 264636 |
| 250 | 78619980 (499, 500) | | | | 21908768, 264692 |
| 251 | 84359489 (501, 502) | | | UNCLASSIFIED | 52645156, 29331822, 29331824, 52644045, 265018, 21906765, 21906768, 265020, 27486261, 27486265, 35695763, 18108376, 264556, 264559, 264565 |
| 252 | 79737756 (503, 504) | Novel Protein sim. GBank gij327166[dbj BAA31651] - (AB014576) KIAA0676 protein [Homo sapiens] | | | 264685, 264687, 264832 |
| 253 | 20443124 (505, 506) | Novel Protein sim. GBank gij3038880[embj CAA18513] - (AL022374) putative ATP-dependent DNA helicase [Streptomyces coelicolor] | | helicase | 264604 |
| 254 | 80027421 (507, 508) | Novel Protein sim. GBank gij3915488[sp O34961 YJMB_BACSU - HYPOTHETICAL SYMPORTER IN COTT-RAPA INTERGENIC REGION | | UNCLASSIFIED | 264508, 264906, 264602, 264687, 265021, 264488 |
| 255 | 11388315 (508, 510) | Novel Protein sim. GBank gij1665720[dbj BAA04134] - (D17312) diarrheal toxin [Bacillus cereus] | | UNCLASSIFIED | 264593 |
| 256 | 80028158 (511, 512) | Novel Protein sim. GBank gij465787[sp P34422 YL31_CAEEL - HYPOTHETICAL 86.0 KD PROTEIN F44B9.1 IN CHROMOSOME III | Contains protein domain (PF00326) - Prolyl oligopeptidase family | peptidase | 264602, 264692 |
| 257 | 20289282 (513, 514) | Novel Protein sim. GBank gij1172039[sp P42315 SCOA_BACSU - PROBABLE SUCCINYL-COA:3-KETOACID-COENZYME A TRANSFERASE SUBUNIT A (SUCCINYL COA:3-OXOACID COA-TRANSFERASE) (OXC1 A) | Contains protein domain (PF01144) - Coenzyme A transferase | transferase | 264605 |
| 258 | 20459464 (515, 516) | Novel Protein sim. GBank gij3127836[embj CAA18902] - (AL023496) hypothetical protein [Streptomyces coelicolor] | | UNCLASSIFIED | 264604 |
| 259 | 78910152 (517, 518) | | | collagen | 264681, 264686, 264692 |
| 260 | 20378437 (519, 520) | | | UNCLASSIFIED | 264692, 264556 |
| 261 | 20265883 (521, 522) | Novel Protein sim. GBank gij123761[sp P24221 HUTH_STRGR - HISTIDINE AMMONIA-LYASE (HISTIDASE) | Contains protein domain (PF00221) - Phenylalanine and histidine ammonia lyases | UNCLASSIFIED | 264600 |
| 262 | 80189317 (523, 524) | | | UNCLASSIFIED | 285017, 264369 |
| 263 | 88095045 (525, 526) | Novel Protein sim. GBank gij3924708[embj CAA84646] - (Z35597) Weak similarity with sea squirt nidogen precursor protein (bias score 71); cDNA EST EMBL:T02069 comes from this gene; cDNA EST EMBL:D76135 comes from this gene; cDNA EST EMBL:D73147 comes from this gene; cDNA EST EMB... | | UNCLASSIFIED | 264488, 264905, 264906, 264907, 264908, 264909, 264512, 264910, 264758, 264596, 264604, 265019, 264805, 264760, 18108351, 264763, 264764, 264288, 264768, 264768, 264769, 264691, 264692, 264693, 264628, 264634, 264635, 264555, 264636, 264638, 264639 |
| 264 | 87370826 (527, 528) | Novel Protein sim. GBank gij3043734[dbj BAA25531] - (AB011177) KIAA0605 protein [Homo sapiens] | Contains protein domain (PF00047) - Immunoglobulin domain | protease | 264259, 264908, 21808754, 265018, 265019, 265020 |

| | | | | | |
|-----|---------------------|---|--|---------------|--|
| 265 | 95355646 (529, 530) | Novel Protein sim. GBank gij4509624[dj][BAA76834.1] - (AB023207) KIAA0980 protein [Homo sapiens] | | kinase | 264488, 35696286, 29331824, 56182181, 35696052, 264508, 264905, 264908, 264907, 66712502, 264908, 264909, 264511, 264512, 264910, 264592, 264595, 264758, 264598, 55811386, 264600, 265017, 264603, 264604, 264605, 264760, 18108351, 264762, 264681, 264764, 264288, 264766, 264768, 264769, 21906765, 21908767, 21908769, 265020, 264691, 33657023, 33657109, 33657182, 264628, 35696423, 35695855, 264630, 264631, 264632, 264634, 264635, 264636, 264555, 264638, 83373044, 56526488, 87168518, 264564, 264566, 264488 |
| 266 | 79588075 (531, 532) | | | UNCLASSIFIED | 264600 |
| 267 | 11362222 (533, 534) | | | UNCLASSIFIED | 264878 |
| 268 | 78909568 (535, 536) | | | UNCLASSIFIED | 264687, 264769, 264689 |
| 269 | 80025810 (537, 538) | | | UNCLASSIFIED | 264602 |
| 270 | 84361144 (539, 540) | Novel Protein sim. GBank gij4507367[ref]NP_003182.1pTARS - threonyl-tRNA synthetase | | UNCLASSIFIED | 264693 |
| 271 | 79552301 (541, 542) | | | UNCLASSIFIED | 264909, 264693 |
| 272 | 9674778 (543, 544) | Novel Protein sim. GBank gij4980738[g]b AAD33331.1 AE00170 - (AE001707) glucose-1-phosphate adenylyltransferase [Thermotoga maritima] | | synthase | 264908 |
| 273 | 12840694 (545, 546) | Novel Protein sim. GBank gij1168224[sp]P44569[5]NTD_HAEIN - PROBABLE 5'-NUCLEOTIDASE PRECURSOR | | UNCLASSIFIED | 264688 |
| 274 | 39524246 (547, 548) | | | | 264564 |
| 275 | 82787041 (549, 550) | Novel Protein sim. GBank gij3253159 (AF005355) - Translation Initiation factor eIF2C [Oryctolagus cuniculus] | | UNCLASSIFIED | 264807, 264908, 264909, 264766, 264768, 264691, 264632, 264636 |
| 276 | 86671073 (551, 552) | Novel Protein sim. GBank gij134920[sp]P21997[SSGP_VOLCA - SULFATED SURFACE GLYCOPROTEIN 185 (SSG 185) | | | 265008, 60432229 |
| 277 | 80079735 (553, 554) | Novel Protein sim. GBank gij128021[sp]P20964[OBG_BACSU - SPO0B-ASSOCIATED GTP-BINDING PROTEIN | | ribosomalprot | 264600, 18108387 |
| 278 | 12866947 (555, 556) | | | UNCLASSIFIED | 264689 |
| 279 | 95292719 (557, 558) | Novel Protein sim. GBank gij79839[pir]S03812 - uvrB protein - Micrococcus luteus | | nuclease | 264508, 264604, 21906764, 264638, 264557, 264404 |
| 280 | 5603617 (559, 560) | | | | 264259 |
| 281 | 80249599 (561, 562) | Novel Protein sim. GBank gij3123160[sp]Q18964[YLN2_CAEEL - HYPOTHETICAL 46.2 KD TRP-ASP REPEATS CONTAINING PROTEIN D2013.2 IN CHROMOSOME II | | | 18108392, 264634, 264555, 264558, 264557, 264558 |
| 282 | 18598682 (563, 564) | | | UNCLASSIFIED | 265019 |
| 283 | 20614211 (565, 566) | | | UNCLASSIFIED | 264555 |

| | | | | | |
|-----|---------------------|--|--|-------------------|--|
| 284 | 91212160 (567, 568) | Novel Protein sim. GBank gi 2429094 (U56632) - acetyl xylan esterase; AxeA [Thermotoga neapolitana] | Contains protein domain (PF00300) - Phosphoglycerate mutase family | UNCLASSIFIED | 35696052, 29331828, 264508, 264905, 264600, 264602, 264605, 264682, 264764, 56181562, 21906764, 18108376, 264636, 264559, 18108387 |
| 285 | 8757940 (569, 570) | Novel Protein sim. GBank gi 2072674 emb CAB08305 - (Z95120) rhlE [Mycobacterium tuberculosis] | Contains protein domain (PF00270) - DEAD/DEAH box helicase | UNCLASSIFIED | 264603 |
| 286 | 80503235 (571, 572) | Novel Protein sim. GBank gi 765323 bbs 157676 - (S74439) silk fibroin heavy chain (C-terminal) [Bombyx mori]; silkworms. Peptide Partial, 633 aa [Bombyx mori] | UNCLASSIFIED | ATPase-associated | 35696052, 264769, 264638 |
| 287 | 12745521 (573, 574) | Novel Protein sim. GBank gi 765323 bbs 157676 - (S74439) silk fibroin heavy chain (C-terminal) [Bombyx mori]; silkworms. Peptide Partial, 633 aa [Bombyx mori] | UNCLASSIFIED | collagen | 264689 |
| 288 | 20756502 (575, 576) | Novel Protein sim. GBank gi 1870009 emb CAB06860 - (Z92539) hypothetical protein Rv1019 [Mycobacterium tuberculosis] | Contains protein domain (PF00440) - Bacterial regulatory proteins, telR family | UNCLASSIFIED | 264557 |
| 289 | 80043804 (577, 578) | Novel Protein sim. GBank gi 1870009 emb CAB06860 - (Z92539) hypothetical protein Rv1019 [Mycobacterium tuberculosis] | Contains protein domain (PF00440) - Bacterial regulatory proteins, telR family | UNCLASSIFIED | 264593, 264600 |
| 290 | 80430175 (579, 580) | Novel Protein sim. GBank gi 2506684 sp P40120 YDCG_ECOLI - 59.4 PROTEIN IN TRG-RIML INTERGENIC REGION PRECURSOR | UNCLASSIFIED | UNCLASSIFIED | 264768 |
| 291 | 20747431 (581, 582) | Novel Protein sim. GBank gi 2506684 sp P40120 YDCG_ECOLI - 59.4 PROTEIN IN TRG-RIML INTERGENIC REGION PRECURSOR | UNCLASSIFIED | UNCLASSIFIED | 264601 |
| 292 | 80052555 (583, 584) | Novel Protein sim. GBank gi 625182 (L39015) - mitochondrial glutamyl-IRNA synthetase [Saccharomyces cerevisiae] | UNCLASSIFIED | UNCLASSIFIED | 284605 |
| 293 | 80082519 (585, 586) | Novel Protein sim. GBank gi 1718065 sp P53528 UVRD_MYCLE - PUTATIVE DNA HELICASE II HOMOLOG | helicase | UNCLASSIFIED | 264809, 264605, 264687, 264689, 264692 |
| 294 | 79830303 (587, 588) | Novel Protein sim. GBank gi 117422 sp P10040 CRB_DROME - CRUMBS PROTEIN PRECURSOR (95F) | Contains protein domain (PF00008) - EGF-like domain | oncogene | 35696052, 264906, 265011, 264628, 55811576 |
| 295 | 79444160 (589, 590) | Novel Protein sim. GBank gi 1181619 db BAA11565 - (D82364) a variant of TSC-22 [Gallus gallus] | UNCLASSIFIED | UNCLASSIFIED | 52844507, 29331822, 264592, 265020, 264639 |
| 296 | 79607076 (591, 592) | Novel Protein sim. GBank gi 3649789 db BAA33403 - (AB012228) SecA [Vibrio alginolyticus] | UNCLASSIFIED | UNCLASSIFIED | 264508 |
| 297 | 79631287 (593, 594) | Novel Protein sim. GBank gi 5689967 emb CAB52004.1 - (AL109663) putative membrane protein [Streptomyces coelicolor A3(2)] | UNCLASSIFIED | UNCLASSIFIED | 264905, 264687, 264638 |
| 298 | 80418898 (595, 596) | Novel Protein sim. GBank gi 5689967 emb CAB52004.1 - (AL109663) putative membrane protein [Streptomyces coelicolor A3(2)] | UNCLASSIFIED | UNCLASSIFIED | 264905, 264691, 264639, 264766 |

| | | | | | |
|-----|---------------------|---|--|--------------|--|
| 299 | 95283298 (597, 598) | Novel Protein sim. GBank gij220637(dbjBAA01477) - (D10627) zinc finger protein [Mus musculus] | Contains protein domain (PF00096) - Zinc finger, C2H2 type | | 264488, 263994, 56994075, 22278997, 22278998, 22278999, 20281099, 29331824, 29331825, 29331826, 80432289, 29331827, 29331828, 264905, 264906, 264907, 264908, 52644045, 264909, 264511, 265008, 264910, 264595, 264596, 264758, 33657084, 87168559, 265018, 265019, 264764, 264288, 264766, 264687, 56181582, 264769, 21906765, 21906768, 21906769, 33657023, 264692, 33657109, 27486281, 18108370, 264628, 264629, 55811576, 35695855, 264631, 264634, 264635, 264636, 264639, 83373044, 18108387, 87168518, 22279000, 22279002, 264565, 264566, 264567, 264602 |
| 300 | 20711340 (599, 600) | Novel Protein sim. GBank gij145922 (M20981) - iron dicitrate transport protein precursor [Escherichia coli] | | UNCLASSIFIED | |
| 301 | 13511332 (601, 602) | Novel Protein sim. GBank gij1174661[spjP44594]TGT_HAEIN - QUEUINE TRNA-RIBOSYL TRANSFERASE (TRNA-GUANINE TRANSGLYCOSYLASE) (GUANINE INSERTION ENZYME) | | transport | 264687 |
| 302 | 9875260 (603, 604) | Novel Protein sim. GBank gij1174661[spjP44594]TGT_HAEIN - QUEUINE TRNA-RIBOSYL TRANSFERASE (TRNA-GUANINE TRANSGLYCOSYLASE) (GUANINE INSERTION ENZYME) | | | 264908 |
| 303 | 79574895 (605, 606) | Novel Protein sim. GBank gij67985[pir]HJNVAV - helicase (EC 3.6.1.-) - Autographa californica nuclear polyhedrosis virus | | helicase | 264689 264602 |
| 304 | 20711344 (607, 608) | Novel Protein sim. GBank gij728867[spjP40602]APG_ARATH - ANTER-SPECIFIC PROLINE-RICH PROTEIN APG PRECURSOR | | | 264763 |
| 305 | 80412520 (609, 610) | Novel Protein sim. GBank gij1657554[gbjAAB18082.1] - (U73857) hypothetical protein [Escherichia coli] | | UNCLASSIFIED | 263978 |
| 306 | 8515876 (611, 612) | Novel Protein sim. GBank gij1657554[gbjAAB18082.1] - (U73857) hypothetical protein [Escherichia coli] | | UNCLASSIFIED | 265010, 21906768, 265020, 18108374, 263977 |
| 307 | 80222801 (613, 614) | Novel Protein sim. GBank gij1710612[spjQ10793]RNH2_MYCTU - PROBABLE RIBONUCLEASE HII (RNASE HII) | Contains protein domain (PF01351) - Ribonuclease HII | nuclease | 264910, 264600, 264605, 264687, 264689, 264638, 18108387 |
| 308 | 80064305 (615, 616) | Novel Protein sim. GBank gij1710612[spjQ10793]RNH2_MYCTU - PROBABLE RIBONUCLEASE HII (RNASE HII) | | | 264769 |
| 309 | 80504138 (617, 618) | Novel Protein sim. GBank gij5420387[embjCAB46679.1] - (AJ243459) proteophosphoglycan [Leishmania major] | | | 264603 264602 |
| 310 | 80053616 (619, 620) | Novel Protein sim. GBank gij1144522 (U34957) - phosphoribosylaminimidazole succinocarboxamide synthase [Mycobacterium tuberculosis] | | synthase | |
| 311 | 11090659 (621, 622) | Novel Protein sim. GBank gij1144522 (U34957) - phosphoribosylaminimidazole succinocarboxamide synthase [Mycobacterium tuberculosis] | | | |
| 312 | 80054347 (623, 624) | | | UNCLASSIFIED | 264566 264603, 264567 |
| 313 | 80046168 (625, 626) | | | | |

| | | | | | |
|-----|---------------------|--|--|------------------|--|
| 314 | 87645112 (627, 628) | Novel Protein sim. GBank gij3661583 (AF092175) - ikaros [Dario refo] | Contains protein domain (PF00320) - GATA zinc finger | dna_rna_bind | 264259, 60432289, 29331828, 264905, 264906, 264908, 264909, 265008, 264910, 60432229, 33657402, 60433438, 33109954, 265011, 265017, 264603, 285018, 264288, 264768, 264692, 35695763, 264628, 264629, 264639, 60170394, 22279002, 264568 |
| 315 | 82356091 (628, 630) | Novel Protein sim. GBank gij1652620(dj)BAA17540(- (D90907) pyridine nucleotide transhydrogenase beta subunit [Synechocystis sp.] | | | 264508, 264600, 264762, 264687, 264768, 52644229, 264769, 264689, 264635, 264636, 264638, 284486 |
| 316 | 78611071 (631, 632) | | | UNCLASSIFIED | 264693 |
| 317 | 20466944 (633, 634) | Novel Protein sim. GBank gij118244(isp)P24176(DAPE, ECOLI - SUCCINYL-DIAMINOPIMELATE DESUCCINYLAASE (SDAP) | | UNCLASSIFIED | 264605 |
| 318 | 94141838 (635, 636) | Novel Protein sim. GBank gij4680229(gb)AAD27583.1(AF118274) DNB-5 [Homo sapiens] | Contains protein domain (PF00526) Dictyostelium (slime mold) repeats | transport | 264908, 264909, 264910, 264593, 264594, 264760, 264288, 264768, 264769, 21908769, 264691, 264693, 264628, 65274791, 264635, 264636, 264638, 83373044, 22279002, 264568 |
| 319 | 17289360 (637, 638) | Novel Protein sim. GBank gij1148693(emb)CAA60220(- (X86489) rbsC [Clostridium perfringens] | | transport | 265018 |
| 320 | 13527675 (638, 640) | Novel Protein sim. GBank gij2611033(isp)O05314(GLGC_MYCTU - GLUCOSE-1-PHOSPHATE ADENYLYLTRANSFERASE (ADP-GLUCOSE SYNTHASE) (ADP-GLUCOSE PYROPHOSPHORYLASE) | | synthase | 264687 |
| 321 | 84134387 (641, 642) | Novel Protein sim. GBank gij1680716 (U68234) - all-trans-retinoic acid 4-hydroxylase [Dario refo] | | cyto450 | 264509, 264906, 264907, 264908, 265009, 264596, 264764, 264628, 264634, 264635, 264638, 264639, 83373044, 264567 |
| 322 | 66489053 (643, 644) | Novel Protein sim. GBank gij1160355 (U33058) - UNC-89 [Caenorhabditis elegans] | | UNCLASSIFIED | 55811150, 264691, 60431528, 55810764 |
| 323 | 94653725 (645, 646) | | | UNCLASSIFIED | 264488, 265009, 264593, 264628, 264635 |
| 324 | 79174383 (647, 648) | | | | 264687 |
| 325 | 78662691 (649, 650) | | | UNCLASSIFIED | 264693 |
| 326 | 28774974 (651, 652) | | | UNCLASSIFIED | 284288, 18108385 |
| 327 | 79776267 (653, 654) | Novel Protein sim. GBank gij451544 (U04267) - proline-rich cell wall protein [Gossypium barbadense] | | | 264488, 264805, 264509, 264910 |
| 328 | 80253202 (655, 656) | | | UNCLASSIFIED | 264592 |
| 329 | 10173821 (657, 658) | | | UNCLASSIFIED | 284510 |
| 330 | 86597767 (659, 660) | | | UNCLASSIFIED | 284259, 264908 |
| 331 | 78754888 (661, 662) | Novel Protein sim. GBank gij14191358 (AF087825) - claudin-7 [Mus musculus] | | transcriptfactor | 264910, 264687, 264689, 264636, 264567 |
| 332 | 80071440 (663, 664) | Novel Protein sim. GBank gij114049(isp)P19480(AHPF_SALTY - ALKYL HYDROPEROXIDE REDUCTASE SUBUNIT F (ALKYL HYDROPEROXIDE REDUCTASE F52A PROTEIN) | | reductase | 35696423, 264636, 264638, 264565 |
| 333 | 13009555 (665, 666) | | | | 264687 |

| | | | | | |
|-----|---------------------|--|---|---------------|--|
| 334 | 80230771 (667, 668) | Novel Protein sim. GBank gjl322228[pir]S32227 - glutamate dehydrogenase (NADP+) (EC 1.4.1.4) - Corynebacterium glutamicum | Contains protein domain (PF00208) - Glutamate/Leucine/Phenylalanine/Va line dehydrogenase | dehydrogenase | 264905, 264600, 264604, 264486 |
| 335 | 80057026 (668, 670) | Novel Protein sim. GBank gjl2193938[emb]CAB09602] - (Z96800) glpQ2 [Mycobacterium tuberculosis] | | esterase | 264907, 264603, 264693, 18108374, 264636, 18108387 |
| 336 | 80414319 (671, 672) | | | UNCLASSIFIED | 265009, 264766, 264686 |
| 337 | 11090829 (673, 674) | | | | 264602 |
| 338 | 95413134 (675, 676) | Novel Protein sim. GBank gjl5454074[ref]NP_006303.1[psMRT - silencing mediator for retinoid and thyroid hormone receptors | Contains protein domain (PF00249) - Myb-like DNA-binding domain | nucL_recp | 264589, 18108397, 22278998, 29331822, 29321099, 29331824, 58182181, 68714117, 29331825, 35696052, 29331828, 264508, 264509, 264905, 264908, 264907, 264908, 264909, 265006, 265008, 264910, 265009, 264758, 55812038, 65274444, 265011, 87168559, 265017, 265018, 265019, 264760, 55811150, 264681, 264762, 18108351, 264682, 264764, 264766, 264685, 264686, 264768, 52644229, 264689, 55811957, 35695917, 264692, 264693, 264628, 18108370, 18108374, 55811576, 35696423, 35695855, 264635, 264555, 264636, 264556, 264637, 264557, 18108380, 264638, 264558, 264639, 18108381, 83373044, 18108385, 87168518, 60432113 |
| 339 | 11398513 (677, 678) | Novel Protein sim. GBank gjl4001713[dbj]BAA35087.1] - (AB015679) DnaK [Porphyromonas gingivalis] | | eph | 264593 |
| 340 | 80504149 (679, 680) | Novel Protein sim. GBank gjl2842699[sp]Q92353]UBPC SCHPO - PUTATIVE UBIQUITIN CARBOXYL-TERMINAL HYDROLASE C6G9.08 (UBIQUITIN THIOLESTERASE) (UBIQUITIN-SPECIFIC PROCESSING PROTEASE) (DEUBIQUITINATING ENZYME) | | ubiquitin | 264905, 265019, 264769, 18108374 |
| 341 | 11075198 (681, 682) | Novel Protein sim. GBank gjl2688580 (AE001166) - conserved hypothetical protein [Borrelia burgdorferi] | Contains protein domain (PF00290) - Tryptophan synthase alpha chain | isomerase | 264605 |
| 342 | 80054186 (683, 684) | Novel Protein sim. GBank gjl1684738[emb]CAA70601] - (Y09452) Yed J hypothetical protein [Pseudomonas syringae] | | | 264603, 264604 |
| 343 | 20468792 (685, 686) | Novel Protein sim. GBank gjl2117275[emb]CAB09104] - (Z95618) hypothetical protein Rv0807 [Mycobacterium tuberculosis] | | | 264605 |
| 344 | 80428870 (687, 688) | | | UNCLASSIFIED | 264600, 264605, 264768, 18108370, 18108374, 35695855 |
| 345 | 80258853 (689, 690) | Novel Protein sim. GBank gjl3023317[sp]Q48935]APHA_MYCRA - ACETYL POLYAMINE AMINOHYDROLASE | | histone | 264593 |
| 346 | 79831058 (691, 692) | Novel Protein sim. GBank gjl4239787[emb]CAA75437] - (Y15166) NADP-glutamate dehydrogenase [Pseudomonas aeruginosa] | Contains protein domain (PF00208) - Glutamate/Leucine/Phenylalanine/Va line dehydrogenase | dehydrogenase | 264905 |

| | | | | | |
|-----|---------------------|---|---|---------------|---|
| 347 | 79158195 (693, 694) | Novel Protein sim. GBank gii731675[sp]P38795 YH4_YEAST - HYPOTHETICAL 80.7 KD PROTEIN IN ERG7-NMD2 INTERGENIC REGION | | UNCLASSIFIED | 265006, 265008, 265010, 265018, 263967, 263981 |
| 348 | 80020208 (695, 696) | Novel Protein sim. GBank gii1073610[pi]S47672 - ugpB protein - Escherichia coli | | transport | 264602, 18108351, 18108387 |
| 349 | 17282112 (697, 698) | Novel Protein sim. GBank gii3261599[emb]CAB009171 - (Z77137) hypothetical protein Rv1277 [Mycobacterium tuberculosis] | | nuclease | 265007 265009, 264769, 264689, 18108370 |
| 350 | 80502370 (699, 700) | Novel Protein sim. GBank gii2959367[emb]CAA179211 - (AL022117) hypothetical protein [Schizosaccharomyces pombe] | | glycoprotein | 264769, 264905, 264908 |
| 351 | 80501805 (701, 702) | Novel Protein sim. GBank gii4416302[gb]AAD203071 - (AF105716) copla-type pol polyprotein [Zea mays] | | protease | 264595 |
| 352 | 11611585 (703, 704) | Novel Protein sim. GBank gii1174887[sp]P42873 URE1_STAXY - UREASE ALPHA SUBUNIT (UREA AMIDOHYDROLASE) | Contains protein domain (PF00449) - Urease | UNCLASSIFIED | 264604 |
| 353 | 80061653 (705, 706) | Novel Protein sim. GBank gii115157[sp]P16574 BVGA_BORPE - VIRULENCE FACTORS PUTATIVE POSITIVE TRANSCRIPTION REGULATOR BVGA | | UNCLASSIFIED | 264628 |
| 354 | 56826130 (707, 708) | Novel Protein sim. GBank gii115157[sp]P16574 BVGA_BORPE - VIRULENCE FACTORS PUTATIVE POSITIVE TRANSCRIPTION REGULATOR BVGA | Contains protein domain (PF00072) - Response regulator receiver domain | UNCLASSIFIED | 264909, 264595, 264683, 22279002 |
| 355 | 80046344 (709, 710) | Novel Protein sim. GBank gii115157[sp]P16574 BVGA_BORPE - VIRULENCE FACTORS PUTATIVE POSITIVE TRANSCRIPTION REGULATOR BVGA | | oxidase | 264605 |
| 356 | 80043835 (711, 712) | Novel Protein sim. GBank gii2290990 (AF006000) - Brg1 [Bordetella pertussis] | | UNCLASSIFIED | 264768 |
| 357 | 80070566 (713, 714) | Novel Protein sim. GBank gii497637 (J03939) - cytochrome oxidase d subunit 1 [Escherichia coli] | | UNCLASSIFIED | 264604, 264769 |
| 358 | 37032756 (715, 716) | Novel Protein sim. GBank gii2290990 (AF006000) - Brg1 [Bordetella pertussis] | | UNCLASSIFIED | 264594 |
| 359 | 80501488 (717, 718) | Novel Protein sim. GBank gii3510639 (AF049344) - UDP- GalNAc:polypeptide N-acetylglucosaminyltransferase T5 [Rattus norvegicus] | | transferase | 22278996, 264259, 29331822, 29331824, 264605, 55811857, 265022 |
| 360 | 80026748 (719, 720) | Novel Protein sim. GBank gii113764[sp]P25719 AMY1_ECOLI - ALPHA-AMYLASE PRECURSOR (1,4-ALPHA-D-GLUCAN GLUCANOHYDROLASE) | | amylase | 264688 |
| 361 | 80584075 (721, 722) | Novel Protein sim. GBank gii113764[sp]P25719 AMY1_ECOLI - ALPHA-AMYLASE PRECURSOR (1,4-ALPHA-D-GLUCAN GLUCANOHYDROLASE) | | | |
| 362 | 13089485 (723, 724) | Novel Protein sim. GBank gii113764[sp]P25719 AMY1_ECOLI - ALPHA-AMYLASE PRECURSOR (1,4-ALPHA-D-GLUCAN GLUCANOHYDROLASE) | | | |
| 363 | 79750145 (725, 726) | Novel Protein sim. GBank gii2829819[sp]P95171 NUOK_MYCTU - NADH DEHYDROGENASE I CHAIN K (NADH-UBIQUINONE OXIDOREDUCTASE CHAIN 11) (NUO11) | Contains protein domain (PF00420) - NADH-ubiquinone/plastoquinone oxidoreductase chain 4L | dehydrogenase | 264586 264789, 264602, 264604, 264508, 264762, 264638, 264486 |
| 364 | 82443593 (727, 728) | Novel Protein sim. GBank gii2829819[sp]P95171 NUOK_MYCTU - NADH DEHYDROGENASE I CHAIN K (NADH-UBIQUINONE OXIDOREDUCTASE CHAIN 11) (NUO11) | | | |

| | | | | |
|-----|---------------------|---|--|--|
| 365 | 88040288 (728, 730) | Novel Protein sim. GBank gij4928268[gb AAD33924.1] - (AF144237) LOMP protein [Homo sapiens] | Contains protein domain (PF00412) - LIM domain containing proteins | 264488, 21906766, 21906767, 55811576, 21906769, 29148629, 22278995, 22278996, 265020, 265022, 264634, 264691, 264593, 33657023, 33657402, 264693, 264639, 264594, 29331824, 264758, 18108385, 29331827, 87168559, 265018, 22279000, 265019, 264482, 264761, 264681, 18108351, 265017, 264757 |
| 366 | 81821838 (731, 732) | Novel Protein sim. GBank gij4503843[ref NP_003908.1 pGZAD - UNKNOWN | Contains protein domain (PF01602) - Adaplin N terminal region | 60424179, 65274572, 56182575, 22278994, 56994075, 22278998, 264259, 29331822, 29331824, 56182181, 60424269, 88714117, 29331825, 60432289, 29331826, 29331827, 29331828, 264905, 264828, 56182435, 265008, 264512, 265008, 264591, 55812038, 55811386, 265010, 87168559, 265017, 265018, 264804, 265019, 55811150, 264448, 264369, 264288, 264886, 264768, 56181562, 21906769, 55811957, 35695917, 265022, 60170815, 33657023, 65274620, 18108365, 263967, 33657109, 33657349, 35695763, 264828, 18108376, 55811576, 65274791, 35695855, 56182323, 83373044, 60432113, 264563, 264564, 264567 |
| 368 | 79607265 (735, 736) | Novel Protein sim. GBank gij3913029[sp P94967 ALR_MYCSM - ALANINE RACEMASE | UNCLASSIFIED | 264509, 264508, 264804, 264605, 264636 |
| 369 | 95282917 (737, 738) | Novel Protein sim. GBank gij3249559 (AF018261) - EH domain binding protein Epsin [Rattus norvegicus] | UNCLASSIFIED | 264905, 264592, 264605, 264766, 264691 |
| 370 | 88090366 (739, 740) | Novel Protein sim. GBank gij2995299[emb CAA18328] - (AL022269) putative IRNA delta(2)-isopentenylpyrophosphate transferase [Streptomyces coelicolor] | Contains protein domain (PF01715) - IPP transferase | 264905, 264906, 264510, 264600, 264601, 264602, 264603, 265018, 264604, 264605, 265021, 264692, 264636, 264564 |
| 371 | 95292599 (741, 742) | Novel Protein sim. GBank gij2995299[emb CAA18328] - (AL022269) putative IRNA delta(2)-isopentenylpyrophosphate transferase [Streptomyces coelicolor] | UNCLASSIFIED | 264564 |
| 372 | 80021107 (743, 744) | Novel Protein sim. GBank gij2506393[sp P31576 FIXX_ECOLI - FERREDOXIN LIKE PROTEIN | UNCLASSIFIED | 264909, 264905, 264906 |
| 373 | 79863766 (745, 746) | Novel Protein sim. GBank gij3341640[emb CAA13164] - (AJ231122) z611 [Vibrio cholerae] | UNCLASSIFIED | 264905, 264906 |
| 374 | 79847568 (747, 748) | Novel Protein sim. GBank gij5456934[gb AAD43716.1] - (AF152322) protocadherin gamma A2 [Homo sapiens] | cadherin | 55274572, 264259, 29331826, 56182435, 60433356, 60433438, 264757, 55812038, 264758, 55811957, 264690, 33657023 |
| 375 | 91230181 (749, 750) | Novel Protein sim. GBank gij1805408[dbj BAA08970] - (D50453) homologues to nitrite hydratase region 3'-hypothetical protein P47K of P. chlororaphis [Bacillus subtilis] | UNCLASSIFIED | 264759 |
| 376 | 80505214 (751, 752) | Novel Protein sim. GBank gij1805408[dbj BAA08970] - (D50453) homologues to nitrite hydratase region 3'-hypothetical protein P47K of P. chlororaphis [Bacillus subtilis] | UNCLASSIFIED | 264759 |
| 377 | 10338083 (753, 754) | | | 264906 |

| | | | | | |
|-----|---------------------|---|--|---------------|--|
| 378 | 80056153 (755, 756) | Novel Protein sim. GBank gij1076013 pir A49930 - carb protein homolog - Mycobacterium bovis (strain BCG) (fragment) | Contains protein domain (PF00289) - synthase (CPSase) | UNCLASSIFIED | 265008, 264555 |
| 379 | 80503437 (757, 758) | Novel Protein sim. GBank gij216556 dbj BAA02174 - (D12651) glucose dehydrogenase [Escherichia coli] | Contains protein domain (PF01011) - PQQ enzyme repeat | dehydrogenase | 264769 |
| 380 | 80060937 (759, 760) | Novel Protein sim. GBank gij3327136 dbj BAA31636 - (A8014561) KIAA0661 protein [Homo sapiens] | | UNCLASSIFIED | 264604 |
| 381 | 11769027 (761, 762) | | | | 264684 |
| 382 | 80054377 (763, 764) | | | | 264592 |
| 383 | 83258025 (765, 766) | | | | 264595, 265017, 265021, 264638, 87168518, 22279002 |
| 384 | 95314255 (767, 768) | | | UNCLASSIFIED | 264288, 264766, 263967, 85274791, 35695855, 263981, 83373044, 264567 |
| 385 | 10237678 (769, 770) | | | | 264692 |
| 386 | 79633434 (771, 772) | Novel Protein sim. GBank gij1073458 pir S47810 - probable alcohol dehydrogenase (EC 1.1.1.1) - Escherichia coli | Contains protein domain (PF00465) - iron-containing alcohol dehydrogenases | dehydrogenase | 264806 |
| 387 | 17960637 (773, 774) | Novel Protein sim. GBank gij1460074 emb CAB01049 - (Z77250) hypothetical protein Rv2566 [Mycobacterium tuberculosis] | Contains protein domain (PF01841) - Transglutaminase-like superfamily | UNCLASSIFIED | 264760 |
| 388 | 87741376 (775, 776) | Novel Protein sim. GBank gij4240169 dbj BAA74863.1 - (AB020647) KIAA0840 protein [Homo sapiens] | Contains protein domain (PF00646) - F-box domain. | homeobox | 35696286, 264905, 66712502, 60432229, 264593, 60433358, 264686, 264688, 21906765, 264691, 22279000, 264482 |
| 389 | 79316971 (777, 778) | | | UNCLASSIFIED | 18108394, 22278996, 264630, 264556, 22279002 |
| 390 | 80079949 (779, 780) | Novel Protein sim. GBank gij854065 emb CAA56337 - (X83413) U88 [Human herpesvirus 6] | | UNCLASSIFIED | 264600 |
| 391 | 7657302 (781, 782) | | | | 264482 |
| 392 | 79786056 (783, 784) | | | UNCLASSIFIED | 264908 |
| 393 | 33206031 (785, 786) | Novel Protein sim. GBank gij3378523 emb CAA08867 - (AJ009832) cyclomalodextrinase glucanotransferase [Thermotoga neapolitana] | | synthase | 264602, 21906764 |
| 394 | 10104463 (787, 788) | | | | |
| 395 | 80229010 (789, 790) | | | UNCLASSIFIED | 264683 |
| 396 | 20436224 (791, 792) | Novel Protein sim. GBank gij2677780 (U70327) - unknown [Pareitropus polyactis] | Contains protein domain (PF00047) - immunoglobulin domain | UNCLASSIFIED | 264508, 264563 |
| 397 | 80417014 (793, 794) | Novel Protein sim. GBank gij4507909 ref NP_000388.1 pWAS - Wiskott-Aldrich syndrome (eczema-thrombocytopenia) | | | 264556 |
| 398 | 91230517 (795, 796) | Novel Protein sim. GBank gij1518458 (U45998) - mitochondrial solute carrier [Onchocerca volvulus] | Contains protein domain (PF00153) - Mitochondrial carrier proteins | transport | 265007, 265009, 264508, 264556, 264629, 264766 |
| 399 | | | | | 18108398, 22278995, 22278996, 56994075, 22278999, 264259, 29331824, 29331826, 264905, 264908, 265007, 265008, 265009, 21906754, 33657084, 265017, 264448, 264288, 264766, 21908765, 21908766, 21908767, 265020, 265021, 33657023, 33657109, 264628, 35698423, 35695855, 264952, 18108380, 264567, 18108391 |

| | | | | | |
|-----|---------------------|--|---|--|--|
| 398 | 80055278 (797, 798) | Novel Protein sim. GBank gi 3358091 dbj BAA31995 - (AB015974) glycerol kinase [Pseudomonas tolaasii] | Contains protein domain (PF00370) - Kinase FGGY family of carbohydrate kinases | | 264592, 264595 |
| 400 | 94117490 (799, 800) | Novel Protein sim. GBank gi 728835 sp P39192 ALU5_HUMAN - IIII ALU SUBFAMILY SC WARNING ENTRY IIII | Contains protein domain (PF00560) - cadherin Leucine Rich Repeat | | 18108394, 56182575, 22278995, 22278997, 22278998, 264259, 29331824, 265008, 265007, 265009, 60432229, 33657402, 21908754, 265010, 265017, 265018, 265019, 18108351, 18108357, 21908765, 265021, 265022, 264691, 264692, 33657023, 18108370, 65274791, 264634, 264636, 60170394, 56182323 264594 |
| 401 | 11397481 (801, 802) | Novel Protein sim. GBank gi 4928292 gb AA033527.1 AF13211 - (AF132117) FhuA [Slaphylococcus aureus] | transport | | |
| 402 | 95420294 (803, 804) | Novel Protein sim. GBank gi 5689487 dbj BAA83027.1 - (AB028988) KIAA1075 protein [Homo sapiens] | Contains protein domain (PF00017) - phosphatase Src homology domain 2 | | 65274572, 56182575, 35696286, 22278996, 22278998, 264093, 264259, 29331822, 29331824, 29331825, 29331828, 60432289, 29331827, 29331828, 264908, 264907, 264909, 265008, 264511, 265007, 265008, 264910, 264591, 33657402, 60433356, 60433438, 264596, 21908754, 52644296, 265010, 265011, 87168559, 265017, 265018, 265019, 264681, 18108351, 264682, 264448, 264288, 264684, 264766, 264767, 264686, 21908765, 21908766, 21908767, 21908768, 21908769, 55811957, 265020, 265021, 265022, 264690, 264693, 65274620, 35695763, 264628, 18108370, 264629, 18108379, 35696423, 55811576, 264635, 264636, 264557, 264638, 18108385, 22279002, 264563, 264564, 264565, 264568, 264768, 264632, 264639, 264563 264682 |
| 403 | 80439913 (805, 806) | | UNCLASSIFIED | | |
| 404 | 11809865 (807, 808) | | polymerase | | |
| 405 | 79471280 (809, 810) | Novel Protein sim. GBank gi 2661649 emb CAA15755 - (AL009198) dnaE2 [Mycobacterium tuberculosis] | | | 265009, 264682 |
| 406 | 78634172 (811, 812) | | Contains protein domain (PF00159) - Pancreatic hormone peptides | | 18108357, 264693 |
| 407 | 80478229 (813, 814) | | UNCLASSIFIED | | |
| 408 | 80078958 (815, 816) | | UNCLASSIFIED | | 264769 |
| 409 | 5840527 (817, 818) | Novel Protein sim. GBank gi 3047117 (AF059919) - similar to ATP-dependent RNA helicases [Arabidopsis thaliana] | UNCLASSIFIED helicase | | 264600 264259 |

| | | | | | |
|-----|---------------------|---|---|---------------|---|
| 410 | 95357486 (819, 820) | Novel Protein sim. GBank gi 475016 dbj BAA06184 - (D29801) Unknown [Mus musculus] | | UNCLASSIFIED | 264489, 52646365, 52646842, 56181686, 3596286, 52645080, 29331822, 29331824, 56182181, 29331825, 60424269, 35696052, 33656970, 264508, 264509, 264905, 264906, 264907, 264908, 52644045, 264909, 264510, 265007, 264512, 265008, 264910, 33657402, 264758, 52646317, 55811386, 265010, 265011, 265017, 264604, 265018, 55811150, 264762, 264764, 264766, 264887, 264768, 264769, 52644229, 21906768, 265020, 265021, 264534, 52644150, 264892, 33657023, 65274620, 33657109, 33657182, 27488281, 35695783, 264828, 264629, 60431528, 18108376, 263978, 35696423, 35695855, 264632, 264634, 264635, 264637, 264638, 264558, 264639, 56182323, 264559, 60432113, 22279002, 264563, 264565, 264486 |
| 411 | 80501670 (821, 822) | | | UNCLASSIFIED | 264769 |
| 412 | 80241662 (823, 824) | | | | 264907, 264910, 263973, 22279002 |
| 413 | 11076446 (825, 826) | Novel Protein sim. GBank gi 3261784 emb CAB08997 - (Z95558) hpx [Mycobacterium tuberculosis] | | eph | 264605 |
| 414 | 82050554 (827, 828) | Novel Protein sim. GBank gi 129036 sp P20707 OD01_AZOVI - 2-OXOGLUTARATE DEHYDROGENASE E1 COMPONENT (ALPHA-KETOGLUTARATE DEHYDROGENASE) | | dehydrogenase | 18108374, 264760, 264769, 264602, 264638, 264603, 264909, 264605 |
| 415 | 84453144 (829, 830) | Novel Protein sim. GBank gi 4686350 gb AAD31273.1 AF13202 - (AF132025) rhophilin [Drosophila melanogaster] | | UNCLASSIFIED | 264908, 87168518 |
| 416 | 80402775 (831, 832) | Novel Protein sim. GBank gi 2555172 (AF025543) - ArcC; carbamate kinase [Rhizobium etli] | | kinase | 264488, 264600, 264602, 264764, 264636 |
| 417 | 20153787 (833, 834) | Novel Protein sim. GBank gi 1709171 sp P52311 MTX2_XANOR - MODIFICATION METHYLASE XORII (CYTOSINE-SPECIFIC METHYLTRANSFERASE XORII) (M.XORII) | Contains protein domain (PF00145) - C-5 cytosine-specific DNA methylase | | 264605 |
| 418 | 94125841 (835, 836) | | | UNCLASSIFIED | 264689, 264693 |
| 419 | 95314273 (837, 838) | | | collagen | 264908, 264910, 264764, 264639 |
| 420 | 37036349 (839, 840) | Novel Protein sim. GBank gi 3261659 emb CAB03751 - (Z81368) hypothetical protein Rv2419c [Mycobacterium tuberculosis] | Contains protein domain (PF00300) - Phosphoglycerate mutase family | phosphatase | 264769 |
| 421 | 95292942 (841, 842) | Novel Protein sim. GBank gi 2916942 emb CAA17580 - (AL021999) hypothetical protein Rv0981 [Mycobacterium tuberculosis] | Contains protein domain (PF00072) - Response regulator receiver domain | phosphatase | 264908, 264600, 264601, 264803, 264604, 264760, 264769 |
| 422 | 79471293 (843, 844) | Novel Protein sim. GBank gi 231752 sp Q00767 CH61_STRAL - 60 KD CHAPERONIN 1 (PROTEIN CPN60 1) (GROEL PROTEIN 1) (HSP58) | Contains protein domain (PF00118) - TCP-1/cpn60 chaperonin family | eph | 22278996, 264682, 18108376, 18108387 |
| 423 | 79604948 (845, 846) | | | UNCLASSIFIED | 264509 |

| | | | | | | |
|-----|---------------------|--|---|---------------------|---|--|
| 424 | 7896557 (847, 848) | Novel Protein sim. GBank gi 4826814 ref NP_004977.1 pKTN1 - kinesin receptor | | | 265019 | |
| 425 | 80431450 (849, 850) | Novel Protein sim. GBank gi 1703701 bbs 178462 - KRP5=kinesin-related protein [rats, testes, Peptide Partial, 167 aa] | Contains protein domain (PF00225) - Kinesin motor domain | - struct | 264909, 265007, 55811386, 264768, 55810764 | |
| 426 | 80064522 (851, 852) | Novel Protein sim. GBank | | | 264605, 264559 | |
| 427 | 80057232 (853, 854) | gi 231829 sp P29829 COBN_PSEDE - COBN PROTEIN | | UNCLASSIFIED | 264603, 264636 | |
| 428 | 79487798 (855, 856) | Novel Protein sim. GBank gi 81286 pit S22697 - extensin - Vohox cartier (fragment) | | UNCLASSIFIED | 264683 | |
| 429 | 80091252 (857, 858) | Novel Protein sim. GBank gi 1806154 emb CAB06451 - (Z84395) hypothetical protein Rv0888 [Mycobacterium tuberculosis] | | UNCLASSIFIED | 35698423, 35695763, 35695855, 265017, 264584, 264762 | |
| 430 | 80504192 (859, 860) | Novel Protein sim. GBank gi 1806154 emb CAB06451 - (Z84395) hypothetical protein Rv0888 [Mycobacterium tuberculosis] | | reductase | 264508, 264905, 264509, 264908, 264909, 265008, 264600, 264687, 264769, 264689, 264636, 264638, 18108385, 264488 | |
| 431 | 20624249 (861, 862) | Novel Protein sim. GBank | | | 264568 | |
| 432 | 16525372 (863, 864) | Novel Protein sim. GBank gi 3123552 emb CAA18609 - (AL022578) dJ393P12.2 (hypothetical Proline-rich protein KIAA0269 LIKE) [Homo sapiens] | | UNCLASSIFIED | 265020 | |
| 433 | 81494303 (865, 866) | Novel Protein sim. GBank gi 3123552 emb CAA18609 - (AL022578) dJ393P12.2 (hypothetical Proline-rich protein KIAA0269 LIKE) [Homo sapiens] | | UNCLASSIFIED | 264907, 264908, 264909, 264910, 264592, 264595, 264758, 264604, 264760, 264762, 264763, 264636, 264637, 22279002 | |
| 434 | 94326323 (867, 868) | Novel Protein sim. GBank gi 2485272 sp Q89626 CDX2_HUMAN - HOMEBOX PROTEIN CDX-2 (CAUDAL-TYPE HOMEBOX PROTEIN 2)(CDX-3) | Contains protein domain (PF00169) - PH domain | UNCLASSIFIED | 55812038, 56182181, 56181562, 29331828, 35896052, 55810764, 55811576, 65274791, 35695855, 60432113, 55811150, 264636, 264766 | |
| 435 | 80502738 (869, 870) | Novel Protein sim. GBank gi 114105 sp P08532 ARAH_ECOLI - L-ARABINOSE TRANSPORT SYSTEM PERMEASE PROTEIN ARAH | | transport | 264595, 264769 | |
| 436 | 41085953 (871, 872) | Novel Protein sim. GBank | | UNCLASSIFIED | 265020, 22279002 | |
| 437 | 11399291 (873, 874) | Novel Protein sim. GBank gi 2291232 gb AA865351.1 - (AF016427) Contains similarity to Pfam domain: PF00004 (AAA), Score=268.1, E-value=3.7e-77, N=1 [Caenorhabditis elegans] | | UNCLASSIFIED | 264593 | |
| 438 | 11773835 (875, 876) | Novel Protein sim. GBank gi 3242702 (AC003040) - hypothetical protein [Arabidopsis thaliana] | | UNCLASSIFIED | 264686 | |
| 439 | 80018495 (877, 878) | Novel Protein sim. GBank gi 3242702 (AC003040) - hypothetical protein [Arabidopsis thaliana] | | | 264905, 264600, 264602, 264604 | |
| 440 | 79841062 (879, 880) | Novel Protein sim. GBank gi 2291232 gb AA865351.1 - (AF016427) Contains similarity to Pfam domain: PF00004 (AAA), Score=268.1, E-value=3.7e-77, N=1 [Caenorhabditis elegans] | Contains protein domain (PF00004) - ATPases associated with various cellular activities (AAA) | - ATPase_associated | 35696052, 264905, 264908, 264909, 265011, 35698423 | |
| 441 | 20396935 (881, 882) | Novel Protein sim. GBank gi 5639946 gb AAD45904.1 AF16132 - (AF161328) histidine kinase CstS [Corynebacterium diphtheriae] | | | 264605 | |
| 442 | 85281058 (883, 884) | Novel Protein sim. GBank gi 1184790 (U46068) - von Ebner minor salivary gland protein [Mus musculus] | | UNCLASSIFIED | 29331830, 264909 | |
| 443 | 82456427 (885, 886) | Novel Protein sim. GBank gi 5689893 emb CAB52056.1 - (AL109732) putative ATP-binding RNA helicase [Streptomyces coelicolor A3(2)] | | UNCLASSIFIED | 35696052, 264508, 264906, 264512, 264604, 264762, 264769, 264689, 264636 | |
| 444 | 11395897 (887, 888) | Novel Protein sim. GBank gi 1783249 dbj BAA11726 - (D83026) homologous to citrate-sodium symport (citrate transporters); hypothetical [Bacillus subtilis] | | UNCLASSIFIED | 264591 | |

| | | | | | |
|-----|---------------------|--|--|-------------------|--|
| 445 | 79552709 (889, 890) | Novel Protein sim. GBank gij5531272[embjCAB50897.1] - (AJ243800) WSC4 homologue [Kluyveromyces fragilis] | | UNCLASSIFIED | 264693 |
| 446 | 79810937 (891, 892) | Novel Protein sim. GBank gij538413 (L36315) - zinc finger protein [Mus musculus] | Contains protein domain (PF000096) - Zinc finger, C2H2 type | transcript factor | 264509 |
| 447 | 80438888 (893, 894) | Novel Protein sim. GBank gij1542914[embjCAB02185] - (Z80108) fnt [Mycobacterium tuberculosis] | Contains protein domain (PF00551) - Formyl transferase | dehydrogenase | 264768, 55811576 |
| 448 | 80238110 (895, 896) | Novel Protein sim. GBank gij118794[spP10443]DP3A_ECOLI - DNA POLYMERASE III, ALPHA CHAIN | | polymerase | 264508, 264600, 264603, 264605, 264682, 264769, 18108362, 264634, 18108387 |
| 449 | 20460834 (897, 898) | Novel Protein sim. GBank gij118794[spP10443]DP3A_ECOLI - DNA POLYMERASE III, ALPHA CHAIN | | | 264605, 264559 |
| 450 | 94631210 (899, 900) | Novel Protein sim. GBank gij4589508[dbjBAA76775.1] - (AB023148) KIAA0931 protein [Homo sapiens] | Contains protein domain (PF00481) - Protein phosphatase 2C | phosphatase | 65274572, 22278998, 29331824, 29331826, 264906, 264910, 264592, 52646317, 265017, 21908767, 55811957, 56528486, 22279002 |
| 451 | 21433609 (901, 902) | Novel Protein sim. GBank gij2650614 (AE001104) - conserved hypothetical protein [Archaeoglobus fulgidus] | | UNCLASSIFIED | 264486 |
| 452 | 10267278 (903, 904) | Novel Protein sim. GBank gij2493000[spIQ09450]SCOT_CAEEL - PROBABLE SUCCINYL-COA:3-KETOACID-COENZYME A TRANSFERASE PRECURSOR (3-OXOACID COA-TRANSFERASE) | | UNCLASSIFIED | 264692 |
| 453 | 52560098 (905, 906) | Novel Protein sim. GBank gij4468699[embjCAB38153.1] - (AL035591) putative integral membrane export protein [Streptomyces coelicolor] | | UNCLASSIFIED | 264907, 264600 |
| 454 | 39523922 (907, 908) | Novel Protein sim. GBank gij3411053 (AF034863) - synaptic scaffolding molecule [Rattus norvegicus] | Contains protein domain (PF00595) - PDZ domain (Also known as DHR or GLGF) | transferase | 264603 |
| 455 | 13089692 (909, 910) | Novel Protein sim. GBank gij4508075[refJNP_002733.1]pPRKC - protein kinase C, mu | | UNCLASSIFIED | 264687 |
| 456 | 79563081 (911, 912) | Novel Protein sim. GBank gij113158[spP25516]ACO1_ECOLI - ACONITATE HYDRATASE 1 (CITRATE HYDRO-LYASE 1) (ACONITASE 1) | Contains protein domain (PF00130) - Phorbol esters/diacylglycerol binding domain (C1 domain) | UNCLASSIFIED | 264691 |
| 457 | 79831273 (913, 914) | Novel Protein sim. GBank gij113158[spP25516]ACO1_ECOLI - ACONITATE HYDRATASE 1 (CITRATE HYDRO-LYASE 1) (ACONITASE 1) | | | 264905 |
| 458 | 79561227 (915, 916) | Novel Protein sim. GBank gij113158[spP25516]ACO1_ECOLI - ACONITATE HYDRATASE 1 (CITRATE HYDRO-LYASE 1) (ACONITASE 1) | | kinase | 55812038, 265010, 265018, 264681 |
| 459 | 80567359 (917, 918) | Novel Protein sim. GBank gij113158[spP25516]ACO1_ECOLI - ACONITATE HYDRATASE 1 (CITRATE HYDRO-LYASE 1) (ACONITASE 1) | | kinase | 22278997, 264259, 29331826, 265018, 264448, 264369, 21908765, 35696423 |
| 460 | 79245890 (919, 920) | Novel Protein sim. GBank gij1168574[spP42464]ATPB_CORGL - ATP SYNTHASE BETA CHAIN | | UNCLASSIFIED | 264906 |
| 461 | 95287618 (921, 922) | Novel Protein sim. GBank gij1168574[spP42464]ATPB_CORGL - ATP SYNTHASE BETA CHAIN | | synthase | 264602, 264605, 264768, 264769, 265021, 33857023, 264559 |

| | | | | | |
|-----|---------------------|---|--|--------------|---|
| 462 | 78060589 (923, 924) | Novel Protein sim. GBank gij1348891spIP45597IPTF1_XANCP - MULTIPHOSPHORYL TRANSFER PROTEIN (MTP) (CONTAINS: PHOSHOENOLPYRUVATE-PROTEIN PHOSPHOTRANSFERASE (PHOSPHOTRANSFERASE SYSTEM, ENZYME I); PHOSPHOCARRIER PROTEIN HPR (PROTEIN H); PTS SYSTEM, FRUCTOSE-SPECIFIC IIA COMPONENT ... | Contains protein domain (PF00381) - PEP-utilizing enzymes | UNCLASSIFIED | 264807 |
| 463 | 78786417 (925, 926) | Novel Protein sim. GBank gij854065[emb]CAA58337] - (X63413) U88 [Human herpesvirus 6] | | UNCLASSIFIED | 264905, 264906, 264908, 264909, 264910, 264591, 264595, 265011, 264632, 264635, 264636, 264637, 264638, 264639 264634 |
| 464 | 82340151 (927, 928) | Novel Protein sim. GBank gij588977[emb]CAB52137.1] - (AJ242832) calpain [Homo sapiens] | Contains protein domain (PF00848) - Calpain family cysteine protease | UNCLASSIFIED | 265017, 21906764, 265020 |
| 465 | 83005730 (929, 930) | Novel Protein sim. GBank gij1806175[emb]CAB06470] - (Z84395) rpsC [Mycobacterium tuberculosis] | Contains protein domain (PF00417) - Ribosomal protein S3, N-terminal domain. | UNCLASSIFIED | 264605, 264559 |
| 466 | 20460845 (931, 932) | Novel Protein sim. GBank gij548705spIP36949[RBSB_BACSU - D-RIBOSE-BINDING PROTEIN PRECURSOR | | UNCLASSIFIED | 264764 |
| 468 | 52562208 (935, 936) | Novel Protein sim. GBank gij2114024[emb]CAB08957] - (Z95558) grcC1 [Mycobacterium tuberculosis] | | UNCLASSIFIED | 264692 |
| 469 | 19520527 (937, 938) | Novel Protein sim. GBank gij2909459[emb]CAA17347] - (AL021929) cobQ [Mycobacterium tuberculosis] | | UNCLASSIFIED | 264488 |
| 470 | 80502756 (939, 940) | Novel Protein sim. GBank gij114921spIP17447[BETT_ECOLI - HIGH-AFFINITY CHOLINE TRANSPORT PROTEIN | synthase | | 264602, 264769 |
| 471 | 17937351 (941, 942) | Novel Protein sim. GBank gij114921spIP17447[BETT_ECOLI - HIGH-AFFINITY CHOLINE TRANSPORT PROTEIN | transport | | 265019 |
| 472 | 80047458 (943, 944) | Novel Protein sim. GBank gij862343 (L10908) - Gcap1 gene product [Mus musculus] | | UNCLASSIFIED | 264596, 264685, 264557 264369 |
| 473 | 20558793 (945, 946) | | | UNCLASSIFIED | 22278997, 264692, 264288 |
| 474 | 80593365 (947, 948) | | | UNCLASSIFIED | 264907, 264908, 264511, 265009, 264762, 264448, 264636, 264638 |
| 475 | 82454685 (949, 950) | | | UNCLASSIFIED | 65274572, 60432049, 264259, 264508, 52844045, 55812038, 264758, 265011, 264288, 264686, 52844229, 65274791, 264638, 264586 |
| 476 | 94143857 (951, 952) | Novel Protein sim. GBank gij5453858[ref]NP_006329.1pGAC1 - glioma amplified on chromosome 1 protein (leucine-rich) | glycoprotein | | 264638, 264586 |
| 477 | 79175833 (953, 954) | | | UNCLASSIFIED | 264638 |
| 478 | 79633483 (955, 956) | | | UNCLASSIFIED | 264690, 264693 |
| 479 | 80189746 (957, 958) | | | UNCLASSIFIED | 264686, 35695855, 265008, 264631, 264910, 264632, 264638, 265018, 264369, 264909 |
| 480 | 79390729 (959, 960) | Novel Protein sim. GBank gij1127551 (U18939) - orf2 [Batrachocottus baikalensis] | mapolymerase | | 264369 |
| 481 | 79624578 (961, 962) | | | UNCLASSIFIED | 264693 |
| 482 | 83050611 (963, 964) | Novel Protein sim. GBank gij4063042 (AF068085) - GP800; mucin-like glycoprotein [Cryptosporidium parvum] | | UNCLASSIFIED | 264909, 264686, 264768, 264693, 55811576, 56182323, 18108385 |

| | | | | | |
|-----|---------------------|--|---|--|--|
| 483 | 20283306 (865, 866) | Novel Protein sim. GBank gjl2104303[emb]CAB08632] - (Z95387) hypothetical protein RV2610c [Mycobacterium tuberculosis] | Contains protein domain (PF00534) - Glycosyl transferases group 1 | 284600 | |
| 484 | 11618046 (867, 868) | Novel Protein sim. GBank gjl3450883 (AF083334) - fibroin [Antheraea pernyi] | | UNCLASSIFIED | 284594 |
| 485 | 80191234 (869, 870) | Novel Protein sim. GBank gjl5042272[emb]CAB44526.1] - (AL078618) nuoF. NADH dehydrogenase subunit [Streptomyces coelicolor] | | UNCLASSIFIED | 284369, 21806765, 22279000, 22279002 |
| 486 | 80059042 (871, 872) | | | dehydrogenase | 284604 |
| 487 | 11813339 (873, 874) | | | | 284638 |
| 488 | 91222383 (875, 876) | Novel Protein sim. GBank gjl5724778[gbl]AAC53522.2] - (AF012273) rho-type GTPase-activating protein rhoGAPX-1 [Mus musculus] | Contains protein domain (PF00620) - RhoGAP domain | 284686, 66714117, 284768, 18108385, 55811576, 265008, 265008, 265009, 265019, 22279002, 264259, 18108370, 264907, 284784, 56182323, 264288, 264693 | |
| 489 | 10867710 (877, 878) | Novel Protein sim. GBank gjl3882223[dbj]BAA34471.1] - (AB018294) KIAA0751 protein [Homo sapiens] | | kinase | 264639 |
| 490 | 95361124 (879, 880) | Novel Protein sim. GBank gjl82091[pir]JAZ5494 - hydroxyproline-rich glycoprotein - tomato (fragment) | | collagen | 22278996, 29331822, 29331828, 264107, 264909, 264110, 265009, 264592, 264593, 80433358, 264288, 264693, 263974, 263976, 20281071, 60432113 |
| 491 | 80495412 (881, 882) | Novel Protein sim. GBank gjl2894206[emb]CAA17072] - (AL021840) hypothetical protein RV3258c [Mycobacterium tuberculosis] | | UNCLASSIFIED | 264769 |
| 492 | 87421264 (883, 884) | | | | 264600 |
| 493 | 11692842 (885, 886) | | | UNCLASSIFIED | 264638 |
| 494 | 87726604 (887, 888) | Novel Protein sim. GBank gjl5262605[emb]CAB45743.1] - (AL080150) hypothetical protein [Homo sapiens] | | UNCLASSIFIED | 264489, 35696286, 60432289, 29331828, 35696052, 264509, 264905, 264906, 264907, 264908, 264909, 264510, 264511, 265009, 264910, 33657402, 264762, 264764, 264768, 264769, 264688, 21906765, 21906769, 35695917, 265020, 264693, 33657109, 264629, 35696423, 35695855, 284634, 264638 |
| 495 | 80026599 (889, 890) | Novel Protein sim. GBank gjl2791517[emb]CAA16054] - (AL021246) hypothetical protein RV2477c [Mycobacterium tuberculosis] | Contains protein domain (PF00005) - ABC transporter | transport | 264602, 264682, 264638 |
| 496 | 78985624 (891, 892) | Novel Protein sim. GBank gjl230281[dbj]R669] - 434 Repressor (Amino-Terminal Domain) (R1-69) | Contains protein domain (PF01381) - Helix-turn-helix | | 264601, 265021 |
| 497 | 78949661 (893, 894) | Novel Protein sim. GBank gjl129736[sp]P28225[PDHX_ECOLI - PYRIDOXAMINE 5'-PHOSPHATE OXIDASE (PNP/PMP OXIDASE)] | | oxidase | 265006 |

| | | | | | |
|-----|-----------------------|---|--|--------------|--|
| 498 | 88095488 (995, 996) | Novel Protein sim. GBank gij1145789 (U41662) - neuroigin 2 (Rattus norvegicus) | Contains protein domain (PF00135) Carboxylesterases | esterase | 264259, 29331826, 35696052, 264508, 264509, 264905, 264908, 264907, 264908, 264909, 264510, 264511, 265009, 264910, 264591, 33857402, 264758, 265010, 265011, 264600, 264601, 264605, 264683, 264764, 264768, 264767, 264768, 264687, 264769, 21906767, 33657023, 284693, 264628, 284629, 35696423, 264630, 264632, 264634, 264635, 264637, 264638, 264558, 264639, 18108385, 264563, 264564, 264565, 264566, 264567 |
| 499 | 20438222 (987, 998) | Novel Protein sim. GBank gij97480(pir)S19739 - integral membrane protein - Rhodobacter capsulatus | | UNCLASSIFIED | 264605 |
| 500 | 11076810 (999, 1000) | | | | 264605 |
| 501 | 13418034 (1001, 1002) | Novel Protein sim. GBank gij5708250(emb)CAB52363.1] - (AL109747) putative integral membrane protein [Streptomyces coelicolor A3(2)] | | UNCLASSIFIED | 264688 |
| 502 | 80021176 (1003, 1004) | Novel Protein sim. GBank gij4468678(emb)CAB38132.1] - (AL035591) glucose-6-phosphate isomerase [Streptomyces coelicolor] | Contains protein domain (PF00342) - Phosphoglucose isomerase | isomerase | 22278996, 265011, 264602, 264605, 264635 |
| 503 | 20264483 (1005, 1006) | | | UNCLASSIFIED | 264584 |
| 504 | 10887321 (1007, 1008) | | | UNCLASSIFIED | 264687 |
| 505 | 95003068 (1009, 1010) | | | UNCLASSIFIED | 264369 |
| 506 | 16454292 (1011, 1012) | Novel Protein sim. GBank gij4033509(sp)P02598(CALM_TETPY - CALMODULIN | Contains protein domain (PF00036) - EF hand | struct | 265010 |
| 507 | 20451598 (1013, 1014) | Novel Protein sim. GBank gij2501069(sp)Q46127(SYW_CLOLO - TRYPTOPHANYL-TRNA SYNTHETASE (TRYPTOPHAN--TRNA LIGASE) (TRPRS) | | UNCLASSIFIED | 264604 |
| 508 | 79841424 (1015, 1016) | Novel Protein sim. GBank gij466088(sp)P34618(YO82_CAEEL - HYPOTHETICAL 33.8 KD PROTEIN ZK1236.2 IN CHROMOSOME III | | UNCLASSIFIED | 264908 |
| 509 | 11776386 (1017, 1018) | | | | 264638 |
| 510 | 83373465 (1019, 1020) | | | UNCLASSIFIED | 264687, 264639 |
| 511 | 16525578 (1021, 1022) | | | | 265007 |
| 512 | 20399484 (1023, 1024) | Novel Protein sim. GBank gij2497419(sp)P55635(Y4RB_RHISN - PUTATIVE INTEGRASE/RECOMBINASE Y4RB | | UNCLASSIFIED | 264565 |
| 513 | 79457404 (1025, 1026) | Novel Protein sim. GBank gij1276897 (U41809) - cyclin J [Drosophila melanogaster] | Contains protein domain (PF00134) - Cyclin | cyclin | 264683, 264689, 35696423, 264639 |
| 514 | 79813805 (1027, 1028) | Novel Protein sim. GBank gij1184790 (U46066) - von Ebner minor salivary gland protein [Mus musculus] | | UNCLASSIFIED | 29331830, 264909 |
| 515 | 79462591 (1029, 1030) | | | | 22278999, 264690 |
| 516 | 8862020 (1031, 1032) | Novel Protein sim. GBank gij2127400(pir)S65770 - maltodigosyltrehalose trehalohydrolase - Arthrobacter sp. (strain Q36) | | amylase | 264910 |

| | | | | | |
|-----|-----------------------|---|---|-------------------|--|
| 517 | 95292894 (1033, 1034) | Novel Protein sim. GBank gi 2983605(AE000725) - ribose 5 phosphate isomerase B [Aquifex aeolicus] | | isomerase | 265018, 264605, 264764, 264766, 264687, 264891, 264585 |
| 518 | 8491831 (1035, 1036) | Novel Protein sim. GBank gi 854065[emb]CAA58337] - (X83413) U88 [Human herpesvirus 6] | | UNCLASSIFIED | 264487 |
| 519 | 91677886 (1037, 1038) | Novel Protein sim. GBank gi 5689355[dbj]BAA83073.1] - (AB024075) B120 [Homo sapiens] | Contains protein domain (PF01388) - ARID DNA binding domain | dna_rna_bind | 52644507, 22278997, 22278998, 60432049, 264259, 52645080, 29331824, 66714117, 60424269, 29331826, 35696052, 264905, 29331830, 66712502, 264511, 265007, 264591, 60432229, 33657402, 60433438, 21906754, 33109954, 52644296, 87168474, 87168559, 265017, 265018, 264604, 265019, 264681, 264448, 264369, 264288, 264685, 21906765, 21906766, 21906767, 21906769, 265021, 60170615, 33657023, 264692, 52645129, 33657109, 27486262, 27486264, 35895763, 18108370, 264629, 52644332, 56182323, 264639, 83373044, 18108385, 56526486, 60432113 |
| 520 | 79859188 (1039, 1040) | Novel Protein sim. GBank | | transport | 264769 |
| 521 | 11076821 (1041, 1042) | gi 169126[sp]P46839[CTPA_MYCLE - CATION-TRANSPORTING P-TYPE ATPASE A | | | 264605 |
| 522 | 80435060 (1043, 1044) | Novel Protein sim. GBank | Contains protein domain (PF00294) - pfkB family carbohydrate kinase | kinase | 264905, 264768 |
| 523 | 18356013 (1045, 1046) | Novel Protein sim. GBank gi 2132243[pir]S61028 - hypothetical protein YPL236c - yeast (Saccharomyces cerevisiae) | | UNCLASSIFIED | 264629 |
| 524 | 80261805 (1047, 1048) | Novel Protein sim. GBank gi 4033608[dbj]BAA35136] - (AB012308) B2HC [Anthracidaris crassispina] | | ATPase_associated | 264092, 264596, 265011 |
| 525 | 78610046 (1049, 1050) | Novel Protein sim. GBank gi 4106610[emb]CAA21365] - (AL031866) ORF42, len=386 aa, similarity to an aminotransferase, in P95957 Sulfolobus solfataricus (401 aa), 33.1% identity in 393 aa overlap. FastA scores: opt-488, E(-): 8.5e-24, In Q64602 R. norvegicus,(425 aa), 28.6% ident... | | UNCLASSIFIED | 264907 |
| 526 | 38827630 (1051, 1052) | Novel Protein sim. GBank | | UNCLASSIFIED | 264758 |
| 527 | 80504729 (1053, 1054) | Novel Protein sim. GBank | | UNCLASSIFIED | 264769 |
| 528 | 65484134 (1055, 1056) | Novel Protein sim. GBank | | UNCLASSIFIED | 56182575, 265017, 265018 |
| 529 | 17936810 (1057, 1058) | Novel Protein sim. GBank | | hydrolase | 265019 |
| 530 | 10887336 (1059, 1060) | Novel Protein sim. GBank gi 24215[JXUA_ECOLI - MANNONATE DEHYDRATASE (D-MANNONATE HYDROLASE)] | | UNCLASSIFIED | 264687 |
| 531 | 80226576 (1061, 1062) | Novel Protein sim. GBank gi 42144[emb]CAA25200] - (X00513) NusA protein (nusA) [Escherichia coli] | | UNCLASSIFIED | 264555, 264556, 264557, 264558, 18108385 |
| 532 | 90933444 (1063, 1064) | Novel Protein sim. GBank gi 5262640[emb]CAB45758.1] - (AL080170) hypothetical protein [Homo sapiens] | Contains protein domain (PF00822) - SPRY domain | UNCLASSIFIED | 264488, 264490, 264259, 264592, 264760, 265021, 264690, 263976, 264558 |

| | | | | | |
|-----|-----------------------|--|---|-------------------|--|
| 533 | 87761531 (1065, 1066) | Novel Protein sim. GBank gi1488363[gb AAD31593.1 AF11229] integral inner nuclear membrane protein MAN1 [Homo sapiens] | | | 264907, 264909, 264768, 35695917, 264630, 264555 |
| 534 | 82368264 (1067, 1068) | Novel Protein sim. GBank gi12985352[emb CAA04906.1 - (AJ001206) pep1 [Streptomyces coelicolor] | UNCLASSIFIED | | 264905, 265011, 264601, 264602, 264605, 264762, 264768, 265020, 264693, 264636 264906 |
| 535 | 78641850 (1069, 1070) | Novel Protein sim. GBank gi13878636[emb CAA89531 - (Z49128) similar to cAMP-dependant protein kinase; cDNA EST EMBL: T00719 comes from this gene; cDNA EST yk465d8.3 comes from this gene; cDNA EST yk465d8.5 comes from this gene; cDNA EST yk492f4.3 comes from this gene; cDNA EST y... | Contains protein domain (PF00069) - Eukaryotic protein kinase domain | ATPase-associated | |
| 536 | 79907207 (1071, 1072) | Novel Protein sim. GBank gi12495828[sp P55757 YOH1_SERMA - HYPOTHETICAL 10.1 KD PROTEIN IN BIOA 5/REGION | | reductase | 18108376, 264905, 264908, 264907, 264909 |
| 537 | 94147448 (1073, 1074) | Novel Protein sim. GBank gi1134920[sp P21997 SSGP_VOLCA - SULFATED SURFACE GLYCOPROTEIN 185 (SSG 185) | Contains protein domain (PF00595) - PDZ domain (Also known as DHR or GLGF). | collagen | 265008, 264605, 65274781 29331822, 29331824, 29331825, 29331826, 29331827, 264908, 52644045, 33657402, 265017, 264762, 264683, 264288, 264685, 21906765, 35695763, 264558, 60170394, 264559, 22279002 264602, 265019 |
| 538 | 26396269 (1077, 1078) | Novel Protein sim. GBank gi12498433[sp Q12341 HAT1_YEAST - HISTONE ACETYLTRANSFERASE | | histone | |
| 540 | 79637077 (1078, 1080) | Novel Protein sim. GBank gi13882241[db BAA34480.1 - (AB018303) KIAA0760 protein [Homo sapiens] | Contains protein domain (PF00096) - Zinc finger, C2H2 type | transcriptiador | 264893 18108394, 22278997, 22278998, 264259, 264112, 265009, 33657402, 55812038, 52646317, 265017, 21906765, 264693, 55811576, 264635, 56526486, 264568 264910, 265018, 264689, 264638, 264486 |
| 542 | 85295836 (1083, 1084) | Novel Protein sim. GBank gi15042272[emb CAB44526.1 - (AL078618) nuoF, NADH dehydrogenase subunit [Streptomyces coelicolor] | | dehydrogenase | |
| 543 | 79786280 (1085, 1086) | Novel Protein sim. GBank gi12791398[emb CAA15994 - (AL021184) hypothetical protein Rv1464 [Mycobacterium tuberculosis] | | UNCLASSIFIED | 264602, 264908 |
| 544 | 20437191 (1087, 1088) | Novel Protein sim. GBank gi1350855[sp P19176 RPOC_PSEPU - DNA-DIRECTED RNA POLYMERASE BETA' CHAIN (TRANSCRIPTASE BETA' CHAIN) (RNA POLYMERASE BETA' SUBUNIT) | | UNCLASSIFIED | 264605 |
| 545 | 80434504 (1089, 1090) | Novel Protein sim. GBank gi14887211[gb AAD32237.1 AF14744 - (AF147449) penicillin binding protein 1B [Pseudomonas aeruginosa] | | | 264768, 264634, 264907, 264592, 264809 264600, 264602, 21908765 |
| 546 | 80249016 (1091, 1092) | Novel Protein sim. GBank gi11350855[sp P19176 RPOC_PSEPU - DNA-DIRECTED RNA POLYMERASE BETA' CHAIN (TRANSCRIPTASE BETA' CHAIN) (RNA POLYMERASE BETA' SUBUNIT) | | mapolymerase | 264604 |
| 547 | 11077563 (1093, 1094) | Novel Protein sim. GBank gi12330021 (AF019250) - kinesin- related protein; KRP; Costal2 [Drosophila melanogaster] | | UNCLASSIFIED | 264488, 264905, 264910, 264760, 264693, 264639, 264563, 264564 |

| | | | | | |
|-----|-----------------------|---|---|------------------|--|
| 548 | 95421904 (1097, 1098) | Novel Protein sim. GBank gi 4337460 gb AAD18133 - (AF056195) neuroblastoma-amplified protein [Homo sapiens] | | UNCLASSIFIED | 264488, 65274572, 18108388, 22278995, 22278996, 22278997, 22278998, 22278999, 264259, 29331824, 66714117, 29331825, 28331826, 35896052, 265007, 265008, 264910, 264592, 33657402, 33109954, 265017, 265018, 265019, 18108351, 264448, 264764, 264369, 264288, 264766, 264686, 264688, 21906765, 21906766, 21906767, 21906768, 21906769, 265020, 284691, 33657023, 284692, 264693, 65274620, 52845129, 33657109, 27486261, 27486262, 27486264, 33657349, 55811576, 18108387, 80432113, 22279002 |
| 550 | 10886816 (1099, 1100) | Novel Protein sim. GBank | | UNCLASSIFIED | 264688 |
| 551 | 80439990 (1101, 1102) | Novel Protein sim. GBank gi 3122893 sp P94985 SYFB_MYCTU - PHENYLALANYL-TRNA SYNTHETASE BETA CHAIN (PHENYLALANINE--TRNA LIGASE BETA CHAIN) (PHERS) | | UNCLASSIFIED | 264908, 264909, 264768 |
| 552 | 84672870 (1103, 1104) | Novel Protein sim. GBank gi 552087 (M33753) - crumbs protein [Drosophila melanogaster] | Contains protein domain (PF00008) - EGF-like domain | UNCLASSIFIED | 264689, 264639, 264563 |
| 553 | 80106002 (1105, 1106) | Novel Protein sim. GBank | | UNCLASSIFIED | 55811957, 264628 |
| 554 | 79818379 (1107, 1108) | Novel Protein sim. GBank gi 5019771 gb AAD37857.1 AF13326 - (AF133263) histidine protein kinase-response regulator hybrid protein CvgSY [Pseudomonas syringae pv. syringae] | kinase | | 264908 |
| 555 | 78986347 (1109, 1110) | Novel Protein sim. GBank gi 131515 sp P02908 PTGA_SALTY - PTS SYSTEM, GLUCOSE-SPECIFIC IIA COMPONENT (EIIA-GLC) (GLUCOSE-PERMEASE IIA COMPONENT) (PHOSPHOTRANSFERASE ENZYME II, A COMPONENT) (EIIA-GLC) | Contains protein domain (PF00358) - phosphoenolpyruvate-dependent sugar phosphotransferase system, EIIA 1 | transport | 264762 |
| 556 | 20457127 (1111, 1112) | Novel Protein sim. GBank gi 3914014 sp P96380 IMFD_MYCTU - TRANSCRIPTION-REPAIR COUPLING FACTOR (TRCF) | | transcriptfactor | 264508, 264605, 264559 |
| 557 | 19523405 (1113, 1114) | Novel Protein sim. GBank gi 5042273 emb CAB44527.1 - (AL078618) nuoE, NADH dehydrogenase subunit [Streptomyces coelicolor] | | dehydrogenase | 264488 |
| 558 | 20724429 (1115, 1116) | Novel Protein sim. GBank gi 170933 sp P45331 METE_HAEIN - 5-METHYL TETRAHYDROPTEROYLTRIGLUTAMATE--HOMOCYSTEINE METHYLTRANSFERASE (METHIONINE SYNTHASE, VITAMIN-B12 INDEPENDENT ISOZYME) (COBALAMIN-INDEPENDENT METHIONINE SYNTHASE) | | UNCLASSIFIED | 264602 |
| 559 | 80084353 (1117, 1118) | Novel Protein sim. GBank gi 4980567 gb AAD35173.1 AE001694 iron(III) ABC transporter, permease protein [Thermotoga maritima] | | UNCLASSIFIED | 264634 |

| | | | | | |
|-----|-----------------------|---|---|----------------|---|
| 560 | 80066533 (1119, 1120) | Novel Protein sim. GBank gi 2492595 sp Q53193 Y4TR_RHISN - PROBABLE PEPTIDE ABC TRANSPORTER ATP-BINDING PROTEIN Y4TR | Contains protein domain (PF00005) - ABC transporter | transport | 18108396, 264906, 264602, 264604, 18108374 |
| 561 | 20293187 (1121, 1122) | | | UNCLASSIFIED | 264600 |
| 562 | 11698161 (1123, 1124) | | | UNCLASSIFIED | 264689 |
| 563 | 79761420 (1125, 1126) | Novel Protein sim. GBank gi 4104925 (AF042276) - poly(hydroxycarbonate) granule associated protein GA2 [Pseudomonas putida] | | UNCLASSIFIED | 264910, 264691 |
| 564 | 56716390 (1127, 1128) | Novel Protein sim. GBank gi 2792310 (AF040570) - unknown [Amycolatopsis mediterranei] | dehydrogenase | | 264592 |
| 565 | 56465618 (1129, 1130) | Novel Protein sim. GBank gi 3449284 (dbj BAA32462) - (AB011532) MEGF8 [Rattus norvegicus] | Contains protein domain (PF00008) - EGF-like domain | synthase | 265010 |
| 566 | 94323888 (1131, 1132) | Novel Protein sim. GBank gi 4539568 (emb CAB38487.1) - (AL035636) putative helicase [Streptomyces coelicolor] | | helicase | 264909, 264510, 265008, 284910, 264758, 264600, 264602, 264604, 264605, 264768, 264687, 264689, 35695917, 264693, 65274620, 264488 |
| 567 | 79560955 (1133, 1134) | Novel Protein sim. GBank gi 100506 pir J17455 - Malate dehydrogenase (oxaloacetate-decarboxylating) (NADP+) (EC 1.1.1.40) - Flavaria trithemia (fragment) | Contains protein domain (PF00390) - Malic enzyme | UNCLASSIFIED | 264681, 264691, 264593 |
| 568 | 94681793 (1135, 1136) | Novel Protein sim. GBank gi 3915843 sp O31212 RS2_STRCO - 30S RIBOSOMAL PROTEIN S2 | Contains protein domain (PF00318) - Ribosomal protein S2 | dehydrogenase | 264689 |
| 569 | 38506897 (1137, 1138) | Novel Protein sim. GBank gi 115122 sp P21627 BRAD_PSEAE - HIGH-AFFINITY BRANCHED-CHAIN AMINO ACID TRANSPORT PROTEIN BRAD | | ribosomal prot | 264565 |
| 570 | 78375927 (1139, 1140) | Novel Protein sim. GBank | | UNCLASSIFIED | 18108376, 18108387, 264565 |
| 571 | 79793961 (1141, 1142) | | | transport | 264907, 264909 |
| 572 | 36996838 (1143, 1144) | Novel Protein sim. GBank gi 4539223 (emb CAB39881.1) - (AL049497) putative integral membrane protein [Streptomyces coelicolor] | | UNCLASSIFIED | 264782 |
| 573 | 20715521 (1145, 1146) | | | UNCLASSIFIED | 265007, 264601 |
| 574 | 13521592 (1147, 1148) | Novel Protein sim. GBank | | | 264636 |
| 575 | 13076416 (1149, 1150) | Novel Protein sim. GBank gi 118794 sp P10443 DP3A_ECOLI - DNA POLYMERASE III, ALPHA CHAIN | | polymerase | 264687 |
| 576 | 20482246 (1151, 1152) | Novel Protein sim. GBank gi 5457625 (emb CAB49116.1) - (AJ248283) PAB2227 [Pyrococcus abyssi] | | | 264605 |
| 577 | 66727102 (1153, 1154) | Novel Protein sim. GBank gi 5042274 (emb CAB44528.1) - (AL078618) nuoD, NADH dehydrogenase subunit [Streptomyces coelicolor] | Contains protein domain (PF00346) - Respiratory-chain NADH dehydrogenase, 49 Kd subunit | dehydrogenase | 35696052, 264636 |
| 578 | 11804477 (1155, 1156) | Novel Protein sim. GBank | | | 264638 |
| 579 | 11794723 (1157, 1158) | Novel Protein sim. GBank gi 1723081 sp Q11046 Y089_MYCTU - HYPOTHETICAL ABC TRANSPORTER ATP-BINDING PROTEIN CY50.09 | | transport | 264682, 264556 |

| | | | | | | |
|-----|-----------------------|---|-----------------------------------|---------------|--------------|--|
| 580 | 80059417 (1159, 1160) | | | | | 22278999, 35696052, 284555, 264556, 284558 |
| 581 | 79230833 (1161, 1162) | | | | UNCLASSIFIED | 285008, 264564 |
| 582 | 80049817 (1163, 1164) | Novel Protein sim. GBank gij3243131 (AF045777) - IiIn [Drosophila melanogaster] | Contains protein domain (PF00047) | struct | | 265021, 264555, 264557 |
| 583 | 79321392 (1165, 1166) | Novel Protein sim. GBank gij2501162[sp]P7726[YAJR_ECOLI - HYPOTHETICAL 49.0 KD PROTEIN IN ABPA-CYOE INTERGENIC REGION] | Immunoglobulin domain | transport | | 264594 |
| 584 | 78845024 (1167, 1168) | Novel Protein sim. GBank gij3882221[dbj]BAA34470.1] - (AB018293) KIAA0750 protein [Homo sapiens] | | UNCLASSIFIED | | 264488, 264908, 264768, 284687, 35696423 |
| 585 | 79581454 (1169, 1170) | | | UNCLASSIFIED | | 265018, 284684, 21908769 |
| 586 | 38277488 (1171, 1172) | Novel Protein sim. GBank gij4467250[emb]CAB37575] - (AL035569) probable Glu-IRNA Gln amidotransferase subunit [Streptomyces coelicolor] | | UNCLASSIFIED | hydrolase | 264908, 265007 |
| 587 | 80497359 (1173, 1174) | | | | | 264600, 264602, 264605, 264769, 264690, 264557 |
| 588 | 79557239 (1175, 1176) | Novel Protein sim. GBank gij5689519[dbj]BAA83043.1] - (AB028014) KIAA1091 protein [Homo sapiens] | | UNCLASSIFIED | | 265020, 264692 |
| 589 | 79805828 (1177, 1178) | | | UNCLASSIFIED | | 22278996, 264907, 264909, 264510, 265009, 265010, 264687, 264769, 35695917, 18108376, 264634, 264636, 264638 |
| 590 | 79815629 (1179, 1180) | Novel Protein sim. GBank gij2143293[emb]CAB09390] - (Z95972) rpoB [Mycobacterium tuberculosis] | | UNCLASSIFIED | mapolymerase | 264906, 264909 |
| 591 | 10313540 (1181, 1182) | | | | | 264691 |
| 592 | 13889767 (1183, 1184) | | | MHC | | 263972 |
| 593 | 82348699 (1185, 1186) | Novel Protein sim. GBank gij4511983[gb]AAD21543.1] - (AF088896) electroltransfer ubiquinone oxidoreductase [Zymomonas mobilis] | | dehydrogenase | | 264511, 264762, 264769, 264486 |
| 594 | 20212392 (1187, 1188) | Novel Protein sim. GBank gij1272368 (U51896) - L IgE [Vibrio parahaemolyticus] | | UNCLASSIFIED | | 264605 |
| 595 | 10064064 (1189, 1190) | Novel Protein sim. GBank gij131490[sp]P20966[PTFB_ECOLI - PTS SYSTEM, FRUCTOSE-SPECIFIC IIBC COMPONENT (EIIBC-FRU) (FRUCTOSE-PERMEASE IIBC COMPONENT) (PHOSPHOTRANSFERASE ENZYME II, BC COMPONENT) (EIIF-FRU)] | | | | 264769 |
| 596 | 13085170 (1191, 1192) | | | UNCLASSIFIED | | 264636 |
| 597 | 80259003 (1193, 1194) | | | UNCLASSIFIED | | 264592 |
| 598 | 94140216 (1195, 1196) | | | UNCLASSIFIED | | 264758, 55810764, 264555, 264556, 264637, 83373044 |
| 599 | 20385137 (1197, 1198) | Novel Protein sim. GBank gij125328[sp]P04951[KDSB_ECOLI - 3-DEOXY-MANNO- OCTULONATE CYTIDYLTRANSFERASE (CMP-KOO SYNTHETASE) (CMP-2-KETO-3-DEOXYOCTULONIC ACID SYNTHETASE) (CKS)] | | UNCLASSIFIED | | 264603 |
| 600 | 10357663 (1199, 1200) | | | | | |
| 601 | 79610404 (1201, 1202) | Novel Protein sim. GBank gij12127414[pil]S60064 - hypothetical protein 2 - Corynebacterium glutamicum | | UNCLASSIFIED | | 264906 264510 |

| | | | | | |
|-----|-----------------------|--|--|--------------|--|
| 602 | 79250602 (1203, 1204) | Novel Protein sim. GBank gij3522961[gblAAC34243.1] - (AC004411) putative pto kinase [Arabidopsis thaliana] | Contains protein domain (PF00069) - Eukaryotic protein kinase domain | kinase | 265007 |
| 603 | 11466067 (1205, 1206) | | | UNCLASSIFIED | 264595 |
| 604 | 81675420 (1207, 1208) | | | UNCLASSIFIED | 264758 |
| 605 | 20436657 (1209, 1210) | Novel Protein sim. GBank gij1175322[sp]P44917Y883_HAEIN - HYPOTHETICAL PROTEIN H10883 | | UNCLASSIFIED | 264605 |
| 606 | 80334582 (1211, 1212) | Novel Protein sim. GBank gij5020284[gblAAD38043.1]AF15136 - (AF151363) Cdc42 GTPase-activating protein [Mus musculus] | | UNCLASSIFIED | 264764 |
| 607 | 95361506 (1213, 1214) | Novel Protein sim. GBank gij1188864 (M74027) - mucin [Homo sapiens] | | UNCLASSIFIED | 264508, 264908, 85658542, 264682, 264687, 264689, 264534, 18108376, 35896423, 264638, 264555, 264638 |
| 608 | 11810888 (1215, 1216) | | | UNCLASSIFIED | 264682 |
| 609 | 80084775 (1217, 1218) | Novel Protein sim. GBank gij2498701[sp]P55552[Y4LL_RHISN - HYPOTHETICAL 91.8 KD PROTEIN Y4LL | Contains protein domain (PF00989) - PAS domain | UNCLASSIFIED | 264605 |
| 610 | 79629413 (1219, 1220) | | | | 264692 |
| 611 | 87588205 (1221, 1222) | | | | 264508, 264905, 264907, 264908, 264909, 264511, 264910, 264758, 264604, 264684, 264766, 264688, 264692, 264628, 264635, 264638, 264637, 264558 |
| 612 | 95287851 (1223, 1224) | Novel Protein sim. GBank gij1077366[gbl]CAB07118] - (Z92772) recD [Mycobacterium tuberculosis] | Contains protein domain (PF01443) - nuclease Viral (Superfamily 1) RNA helicase | | 264600, 264601, 264604, 264769, 264558, 264565 |
| 613 | 7523475 (1225, 1226) | | | UNCLASSIFIED | 264369 |
| 614 | 79969348 (1227, 1228) | Novel Protein sim. GBank gij5114231[gblAAD40238.1]AF13670 - (AF136709) histidine kinase YycG [Staphylococcus aureus] | | Kinase | 18108372, 264563 |
| 615 | 38586998 (1229, 1230) | Novel Protein sim. GBank gij1339950[gbl]BAA12741] - (D85230) large subunit of NADH-dependent glutamate synthase [Plectonema boryanum] | | synthase | 264600, 264602, 264629 |
| 616 | 20465331 (1231, 1232) | Novel Protein sim. GBank gij544387[sp]P35673[GALE_ERWAM - UDP-GLUCOSE 4-EPIMERASE (GALACTOWALDENASE) (UDP-GALACTOSE 4-EPIMERASE) | | isomerase | 264605 |
| 617 | 91227222 (1233, 1234) | Novel Protein sim. GBank gij2498097[sp]Q60769[TNP3_MOUSE - TUMOR NECROSIS FACTOR, ALPHA-INDUCED PROTEIN 3 (PUTATIVE DNA BINDING PROTEIN A20) (ZINC FINGER PROTEIN A20) | Contains protein domain (PF00641) - Zn-finger in Ran binding protein and others. | Int | 52645156, 21906765, 35698423, 21906768, 21908769, 22278994, 35698286, 22278998, 265020, 265021, 265007, 265008, 264638, 52844150, 33657023, 264692, 264693, 29331822, 29331824, 55812038, 83373044, 56182181, 60424269, 66714117, 29331825, 33657109, 29331826, 33657182, 29331827, 35696052, 29331828, 27486262, 33657349, 56526486, 265018, 265019, 22279002, 264482, 264448, 29331830, 66712502, 264909 |

| | | | | | |
|-----|-----------------------|---|---|--------------|--|
| 618 | 20632843 (1235, 1236) | Novel Protein sim. GBank gij5459388[embjCAB50746.1] - (AL098839) putative aminotransferase [Streptomyces coelicolor] | | isomerase | 264603 |
| 619 | 91227224 (1237, 1238) | | | | 56994075, 29331826, 33656970, 265008, 33857402, 33109954, 87168559, 264448, 18108374, 83373044 |
| 620 | 81183143 (1239, 1240) | Novel Protein sim. GBank gij464335[spjQ05922]DUS2_MOUSE - DUAL SPECIFICITY PROTEIN PHOSPHATASE 2 (DUAL SPECIFICITY PROTEIN PHOSPHATASE PAC-1) | | phosphatase | 28146498, 264758, 264369, 29148627 |
| 621 | 30239251 (1241, 1242) | | | UNCLASSIFIED | 264556, 264558, 264639 |
| 622 | 20456427 (1243, 1244) | Novel Protein sim. GBank gij2633557[embjCAB13060] - (Z99110) yjdf [Bacillus subtilis] | | UNCLASSIFIED | 264605 |
| 623 | 10131788 (1245, 1246) | Novel Protein sim. GBank gij1857710[gbjAAB48482] - (U87224) contactin associated protein [Rattus norvegicus] | Contains protein domain (PF00054) - Laminin G domain | laminin | 264906 |
| 624 | 18534127 (1247, 1248) | Novel Protein sim. GBank gij1705703[spjP52225]CCME_PSEFL - CYTOCHROME C-TYPE BIOGENESIS PROTEIN CYCK | | cytochrome | 264596 |
| 625 | 13084618 (1249, 1250) | Novel Protein sim. GBank gij2894252[embjCAA17114.1] - (AL021841) hypothetical protein Rv3342 [Mycobacterium tuberculosis] | | UNCLASSIFIED | 264688 |
| 626 | 88062603 (1251, 1252) | Novel Protein sim. GBank gij416592[spjP32323]AGA1_YEAST - A-AGGLUTININ ATTACHMENT SUBUNIT PRECURSOR | | UNCLASSIFIED | 29331822, 264905, 264908, 33657023, 33657109, 264558 |
| 627 | 80255457 (1253, 1254) | Novel Protein sim. GBank gij3098418 (AF040944) - P140 [Mus musculus] | | UNCLASSIFIED | 18108394, 264112, 264593, 265022, 264635 |
| 628 | 80077096 (1255, 1256) | Novel Protein sim. GBank gij1711543[spjP50526]SSP1_SCHPO - SERINE/THREONINE-PROTEIN KINASE SSP1 | Contains protein domain (PF00069) - Kinase Eukaryotic protein kinase domain | kinase | 264600 |
| 629 | 79851602 (1257, 1258) | Novel Protein sim. GBank gij1143204 (U34305) - ORF2; Method: conceptual translation supplied by author. [Shigella sonnei] | | isomerase | 264906, 264907 |
| 630 | 39565156 (1259, 1260) | Novel Protein sim. GBank gij3236368 (AF064748) - S3-12 [Mus musculus] | | UNCLASSIFIED | 264490 |
| 631 | 20598718 (1261, 1262) | Novel Protein sim. GBank gij140687[spjP11666]YGGG_ECOLI - HYPOTHETICAL 30.9 KD PROTEIN IN SBM-FBA INTERGENIC REGION (ORF 4) (F286) | | | 263978 |
| 632 | 27843890 (1263, 1264) | | | UNCLASSIFIED | 264906, 264600, 264605, 264769, 264689, 264486 |
| 633 | 80477772 (1265, 1266) | | | UNCLASSIFIED | 264769 |
| 634 | 17938808 (1267, 1268) | | | | 265019 |
| 635 | 79574508 (1269, 1270) | | | UNCLASSIFIED | 264689 |
| 636 | 79810381 (1271, 1272) | | | UNCLASSIFIED | 264596, 264762, 264693 |

| | | | | | |
|-----|-----------------------|---|---|------------------|--|
| 637 | 82455796 (1273, 1274) | Novel Protein sim. GBank gij2326739[embjCAB10953] - (Z98268) recN [Mycobacterium tuberculosis] | | nuclease | 264906, 264907, 264510, 264511, 264601, 264602, 264603, 264604, 264605, 18108351, 264762, 264766, 264687, 264769, 264689, 35695917, 264693, 264634, 264638, 264639, 264559, 18108385 |
| 638 | 14997457 (1275, 1276) | Novel Protein sim. GBank gij4676682[embjCAB41074.1] - (AL049645) putative large ATP-binding protein [Streptomyces coelicolor] | | | 264636 |
| 639 | 80204210 (1277, 1278) | Novel Protein sim. GBank gij4588628[jb BAA76836.1] - (AB023209) KIAA0992 protein [Homo sapiens] | | strut | 264112, 263974 |
| 640 | 17929579 (1279, 1280) | Novel Protein sim. GBank gij1432083 (U00981) - homolog to Skp1p, an evolutionarily conserved kinetochore protein in budding yeast [Arabidopsis thaliana] | Contains protein domain (PF01466) - Skp1 family | mapolymerase | 265009, 265010 |
| 641 | 79636398 (1281, 1282) | | | UNCLASSIFIED | 264693 |
| 642 | 19898737 (1283, 1284) | | | UNCLASSIFIED | 264565 |
| 643 | 81516220 (1285, 1286) | | | UNCLASSIFIED | 264635, 264639, 264564 |
| 644 | 11751387 (1287, 1288) | | | UNCLASSIFIED | 264694 |
| 645 | 95010907 (1289, 1290) | | | UNCLASSIFIED | 264906, 264762, 264693, 264639, 264559 |
| 646 | 80069083 (1291, 1292) | | | | 264595, 264566 |
| 647 | 80257085 (1293, 1294) | | | | 264909, 264591 |
| 648 | 80077428 (1295, 1296) | Novel Protein sim. GBank gij4507613[ref NP_003738.1 pTNKS - TANKYRASE | Contains protein domain (PF00023) - Ank repeat | transcriptfactor | 264600 |
| 649 | 80247447 (1297, 1298) | Novel Protein sim. GBank gij1044963[jbbs 169648 - protamine [Monodonta turbinata, gonads, Peptide, 106 aa] | | UNCLASSIFIED | 263978 |
| 650 | 11798316 (1299, 1300) | | | UNCLASSIFIED | 264686 |
| 651 | 11776932 (1301, 1302) | Novel Protein sim. GBank gij1346916[sp P12283 PURA_ECOLI - ADENYLOSUCCINATE SYNTHETASE (IMP--ASPARTATE LIGASE)] | | | 264602, 264638 |
| 652 | 85516704 (1303, 1304) | | | UNCLASSIFIED | 264905, 264907, 264909, 263978, 264637 |
| 653 | 82124947 (1305, 1306) | Novel Protein sim. GBank gij1722977[sp Q10638 Y03C_MYCTU - HYPOTHETICAL 82.8 KD PROTEIN CY130.12C | | UNCLASSIFIED | 22278996, 264510, 264511, 264512, 264593, 21908754, 264603, 264760, 18108376, 264556 |
| 654 | 95010589 (1307, 1308) | | | UNCLASSIFIED | 264905, 264595, 264632 |
| 655 | 79320692 (1309, 1310) | Novel Protein sim. GBank gij130327[sp P26647 PLSC_ECOLI - 1-ACYL-SN-GLYCEROL-3-PHOSPHATE ACYLTRANSFERASE (1-AGP ACYLTRANSFERASE) (1-AGPAT) (LYSOPHOSPHATIDIC ACID ACYLTRANSFERASE) (LPAAT)] | Contains protein domain (PF01553) - Acyltransferase | transferase | 264592 |
| 656 | 80416739 (1311, 1312) | | | UNCLASSIFIED | 264602, 264605, 264766, 264691 |
| 657 | 20811010 (1313, 1314) | | | UNCLASSIFIED | 264557, 264558 |

| | | | | | |
|-----|-----------------------|--|--|------------------|--|
| 658 | 87761815 (1315, 1316) | Novel Protein sim. GBank gij5689493jdbj BAA83030.1 - (AB029001) KIAA1078 protein (Homo sapiens) | | UNCLASSIFIED | 22278996, 60432049, 29331822, 29331824, 29331828, 265007, 265009, 33657402, 33657084, 265017, 264448, 21906765, 21906766, 263967, 20281148, 18108370, 18108374, 264482 |
| 659 | 87718663 (1317, 1318) | Novel Protein sim. GBank gij2137872jirj 48724 - zinc finger protein PFZ - mouse | Contains protein domain (PF00086) - Zinc finger, C2H2 type | transcriptfactor | 22278999, 60432049, 66714117, 29331827, 265007, 264766, 58181562, 18108359, 18108365, 18108370, 18108381 |
| 660 | 81897822 (1319, 1320) | | | UNCLASSIFIED | 264757 |
| 661 | 80028023 (1321, 1322) | Novel Protein sim. GBank gij134180j spIP15401 SACY_BACSU - LEVANSUCRASE AND SUCRASE SYNTHESIS OPERON ANTITERMINATOR | Contains protein domain (PF00874) - Transcriptional antiterminator bglG family | UNCLASSIFIED | 264510, 265009, 264600, 264602, 264603, 264604, 264605, 32833986, 18108376, 264636, 18108387, 22279000 |
| 662 | 20463731 (1323, 1324) | Novel Protein sim. GBank gij454528j gb AAD22450.1 AF11618 - (AF11618) SecA homolog [Adinobacillus actinomycetemcomitans] | | UNCLASSIFIED | 264605 |
| 663 | 20628080 (1325, 1326) | Novel Protein sim. GBank gij5689250j b BAA82881.1 - (AB024335) similar to orf5 [Comamonas testosteroni] | | dehydrogenase | 264605 |
| 664 | 80508512 (1327, 1328) | Novel Protein sim. GBank gij1652848j b BAA17766j - (D90909) DNA photolyase [Synecchocystis sp.] | | UNCLASSIFIED | 264769 |
| 665 | 80078053 (1329, 1330) | Novel Protein sim. GBank gij116941j spIP21640 COBJ_PSEDE - PRECORRIN-3B C17 METHYLTRANSFERASE (PRECORRIN-3 METHYLTRANSFERASE) | | isomerase | 264600 |
| 666 | 78603142 (1331, 1332) | Novel Protein sim. GBank gij3261829j emb CAB10927j - (Z98260) hypothetical protein Rv1230c [Mycobacterium tuberculosis] | | glycoprotein | 264907, 265007 |
| 667 | 94631802 (1333, 1334) | Novel Protein sim. GBank gij568885j dbj BAA82702.1 - (AB017438) Orf5 [Streptomyces coelicolor] | | UNCLASSIFIED | 264689, 264602, 264593 |
| 668 | 82051891 (1335, 1336) | Novel Protein sim. GBank gij3581853j emb CAA20809j - (AL031541) 50S ribosomal protein L20 [Streptomyces coelicolor] | Contains protein domain (PF00453) - Ribosomal protein L20 | ribosomalprot | 264905, 264906, 264908, 264600, 264601, 264603, 264605, 264760, 264689, 264636, 264638, 264639 |
| 669 | 12867154 (1337, 1338) | | | UNCLASSIFIED | 264637 |
| 670 | 80238549 (1339, 1340) | Novel Protein sim. GBank gij2582531 (AF026444) - 2-Isopropylmalate synthase [Streptomyces coelicolor] | | synthase | 264905, 264906, 264908, 264601, 264762, 264766, 264689, 264638, 18108385, 264486 |
| 671 | 78601368 (1341, 1342) | | Contains protein domain (PF00023) - Ank repeat | UNCLASSIFIED | 264690, 264692, 264693, 264636, 18108387 |
| 672 | 79834371 (1343, 1344) | Novel Protein sim. GBank gij2114430 (U92703) - Oil-1/EBF-like-3 transcription factor [Mus musculus] | | transcriptfactor | 264910, 265017 |
| 673 | 82285788 (1345, 1346) | Novel Protein sim. GBank gij4589285j gb AAD26430.1 AF13515 - (AF135154) ferric alkaligin siderophore receptor [Bordetella pertussis] | | | 264759 |
| 674 | 78189259 (1347, 1348) | | | UNCLASSIFIED | 264629 |

| | | | | | |
|-----|-----------------------|---|--|-------------------|---|
| 675 | 87895870 (1349, 1350) | Novel Protein sim. GBank gi 4980755 gb AAD35347.1 AE001708 - (AE001708) D- alanine-D-alanine ligase [Thermotoga maritima] | Contains protein domain (PF01820) - D-ala D-ala ligase | UNCLASSIFIED | 264488, 22278999, 66714117, 264508, 264511, 265008, 60433438, 264600, 264601, 264602, 264603, 264604, 264605, 264762, 264687, 264769, 60431602, 18108374, 264636, 264638 265010 |
| 676 | 78899807 (1351, 1352) | Novel Protein sim. GBank gi 1723568 sp Q10479 YDF7_SCHPO - PUTATIVE GLUCOSYLTRANSFERASE C17C9.07 | | | |
| 677 | 21644312 (1353, 1354) | Novel Protein sim. GBank gi 887208 (U03976) - dynein heavy chain isotype 5C [Tripleneustes grallia] | | ATPase-associated | 264591, 264632 |
| 678 | 84225200 (1355, 1356) | Novel Protein sim. GBank gi 1586274 pr j2203365A - laminin alpha5 [Mus musculus] | Contains protein domain (PF00053) - Laminin EGF-like (Domains III and V) | laminin | 264758, 264682, 264557 |
| 679 | 78868855 (1357, 1358) | Novel Protein sim. GBank gi 3928723 emb CAA22219 - (AL034355) putative ABC transporter [Streptomyces coelicolor] | | UNCLASSIFIED | 22278996, 264693 |
| 680 | 20726424 (1359, 1360) | Novel Protein sim. GBank | | | |
| 681 | 94322017 (1361, 1362) | gi 5174493 ref NP_006080.1 pLAMC - laminin, gamma 3 | Contains protein domain (PF00053) - Laminin EGF-like (Domains III and V) | laminin | 264600, 264602 264102, 264907, 264908, 265006, 264693, 263972, 83373044, 264568 |
| 682 | 11392478 (1363, 1364) | Novel Protein sim. GBank | | UNCLASSIFIED | 264595 |
| 683 | 80083660 (1365, 1366) | gi 4758208 ref NP_004081.1 pDUSP - dual specificity phosphatase 3 (vaccinia virus phosphatase VH1-related) | Contains protein domain (PF00782) - Dual specificity phosphatase, catalytic domain | phosphatase | 264634 |
| 684 | 20465367 (1367, 1368) | Novel Protein sim. GBank gi 5420387 emb CAB46679.1 - (AJ243459) proteophosphoglycan [Leishmania major] | | | 264605 |
| 685 | 80246735 (1369, 1370) | | | | |
| 686 | 79208808 (1371, 1372) | | | | 264909, 263967, 263981 |
| 687 | 80085629 (1373, 1374) | | | | 264631 264693, 264635 |
| 688 | 79853412 (1375, 1376) | Novel Protein sim. GBank gi 2688962 (AF027768) - LspA (Serratia marcescens) | | peptidase | 264907, 264638 |
| 689 | 88064256 (1377, 1378) | Novel Protein sim. GBank gi 3046931 (AF049330) - PPAR gamma coactivator [Mus musculus] | Contains protein domain (PF00076) - RNA recognition motif (a.k.a. RRM, RBD, or RNP domain) | UNCLASSIFIED | 264906, 264907, 265007, 265009, 60433438, 21908754, 264760, 18108358, 21908768, 21906769, 265021, 18108361, 263974, 18108379, 264557, 18108385, 22279002 264510, 264511, 264764, 264769 |
| 690 | 80389750 (1379, 1380) | Novel Protein sim. GBank gi 2498941 sp Q15426 SP62_HUMAN - SPLICEOSOME ASSOCIATED PROTEIN 62 (SAP 62) (SF3A66) | | UNCLASSIFIED | |
| 691 | 81854392 (1381, 1382) | Novel Protein sim. GBank gi 5420387 emb CAB46678.1 - (AJ243459) proteophosphoglycan [Leishmania major] | Contains protein domain (PF00097) - Zinc finger, C3HC4 type (RING finger) | UNCLASSIFIED | 264757 55812038, 55811957, 265018, 55811150, 18108351, 264908, 60431528, 264594 |
| 692 | 83608938 (1383, 1384) | | | | |
| 693 | 79586116 (1385, 1386) | Novel Protein sim. GBank gi 854085 emb CAA58337 - (X83413) U88 [Human herpesvirus 6] | | UNCLASSIFIED | 264635 |
| 694 | 82455983 (1387, 1388) | Novel Protein sim. GBank gi 267327 sp Q01033 VC48_HSVSA - HYPOTHETICAL GENE 48 PROTEIN | | | 22278996, 264510, 264602, 264603, 264762, 264687, 264769, 264693 |

| | | | | | |
|-----|-----------------------|--|--|--------------|--|
| 695 | 94147849 (1389, 1390) | Novel Protein sim. GBank gij4468339[emb]CAB38059.1] - (AJ010901) MUC4 [Homo sapiens] | Contains protein domain (PF000094) - von Willebrand factor type D domain | UNCLASSIFIED | 56182575, 264509, 264905, 264907, 29331830, 264908, 264909, 264511, 265007, 264910, 264758, 264764, 264288, 65274791, 264905, 264595 |
| 696 | 79830882 (1391, 1392) | Novel Protein sim. GBank gij2649950 (AE001058) - glutamine ABC transporter, ATP-binding protein (gho) [Archaeoglobus fulgidus] | Contains protein domain (PF000005) - ABC transporter | transport | |
| 697 | 11767889 (1393, 1394) | Novel Protein sim. GBank gij1731343[sp]Q10694[YY25_MYCTU - HYPOTHETICAL 24.4 KD PROTEIN CY49.25] | Contains protein domain (PF01836) - Transposase | UNCLASSIFIED | 264682 |
| 698 | 66895862 (1395, 1398) | | | UNCLASSIFIED | 264688, 35695917 |
| 699 | 79582558 (1397, 1398) | | | UNCLASSIFIED | 264682 |
| 700 | 79830998 (1399, 1400) | | | UNCLASSIFIED | 264693 |
| 701 | 80230242 (1401, 1402) | Novel Protein sim. GBank gij1001236[db]BAA10477] - (DB4003) hypothetical protein [Synechocystis sp.] | | UNCLASSIFIED | 264488, 264510, 264511, 264602, 264605, 264689 |
| 702 | 79814789 (1403, 1404) | Novel Protein sim. GBank gij2488935[sp]Q46338[SOXG_CORSP - SARCOSINE OXIDASE GAMMA SUBUNIT] | | UNCLASSIFIED | 264909 |
| 703 | 20446820 (1405, 1406) | | | oxidase | 264604 |
| 704 | 94312224 (1407, 1408) | Novel Protein sim. GBank gij3150513 (AF067219) - contains similarity to the kelch/MIPP family [Caenorhabditis elegans] | Contains protein domain (PF01344) - Kelch motif | UNCLASSIFIED | 264288, 56181562, 33657109, 264629, 55811576 |
| 705 | 17632141 (1408, 1410) | Novel Protein sim. GBank gij421091[pir]S30730 - hypothetical protein o208 - Escherichia coli | | UNCLASSIFIED | 265006 |
| 706 | 20288062 (1411, 1412) | Novel Protein sim. GBank gij3024872[sp]Q55790[Y074_SYNY3 - HYPOTHETICAL 52.8 KD PROTEIN SLR0074] | | | 264600 |
| 707 | 20638065 (1413, 1414) | Novel Protein sim. GBank gij3420608[sp]AAC31907.1] - (AF075709) ABC transporter ATP-binding subunit [Pseudomonas putida] | | transport | 264603 |
| 708 | 20708292 (1415, 1416) | Novel Protein sim. GBank gij3649741[emb]CAA03985] - (AJ000281) mucin [Homo sapiens] | | struct | 264601, 264692 |
| 709 | 88001439 (1417, 1418) | Novel Protein sim. GBank gij3080425[emb]CAA18744.1] - (AL022604) putative protein [Arabidopsis thaliana] | | | 18108398, 264637, 264908, 264909 |
| 710 | 11356683 (1419, 1420) | | | | 264369 |
| 711 | 17931418 (1421, 1422) | Novel Protein sim. GBank gij4758686[ref]NP_002323.1[lpLRP1 - low density lipoprotein related protein 1 (alpha-2-macroglobulin receptor)] | Contains protein domain (PF00056) - apolipoprotein repeat class B | | 265019 |
| 712 | 80258164 (1423, 1424) | | | | 264591 |
| 713 | 79263126 (1425, 1426) | Novel Protein sim. GBank gij1703286[sp]Q11056[AM12_MYCTU - PUTATIVE AMIDASE CY50.19C] | | hydrolase | 264906, 264907 |
| 714 | 27847651 (1427, 1428) | Novel Protein sim. GBank gij4502351[ref]NP_001692.1[lpBAAT - bile acid Coenzyme A: amino acid N-acyltransferase; glycine N-choyloltransferase] | | | 264508, 264555 |

| | | | | | | |
|-----|-----------------------|--|---|---------------|--|--|
| 715 | 78639423 (1428, 1430) | Novel Protein sim. GBank gjl1780035 (AE000352) - orf, hypothetical protein [Escherichia coli] | | UNCLASSIFIED | 264907 | |
| 716 | 78559072 (1431, 1432) | | | | 264692 | |
| 717 | 79491842 (1433, 1434) | Novel Protein sim. GBank gjl2494074 (spjP55653) GABD_RHISN - PROBABLE SUCCINATE-SEMIALDEHYDE DEHYDROGENASE (NADP+)(SSDH) | | dehydrogenase | 264636 | |
| 718 | 94319658 (1435, 1436) | Novel Protein sim. GBank gjl3873679 (embjCAA94886) - (Z71178) similar to pro-collagen domains: cDNA EST. EMBL:D27978 comes from this gene: cDNA EST. EMBL:D27977 comes from this gene: cDNA EST. EMBL:D34199 comes from this gene: cDNA EST. EMBL:D64392 comes from this gene: cDNA EST. EMBL... Novel Protein sim. GBank gjl2104302 (embjCAB08631) - (Z95387) hypothetical protein Rv2611c [Mycobacterium tuberculosis] | Contains protein domain (PF00093) - von Willebrand factor type C domain | kinase | 18108392, 22278994, 22278998, 265008, 265018, 264681, 18108354, 264684, 264685, 264686, 264687, 264689, 21906769, 18108361, 264691, 264692, 55810764, 264635, 18108381, 18108382, 83373044, 18108388 | |
| 719 | 17679564 (1437, 1438) | | | UNCLASSIFIED | 265011 | |
| 720 | 79841684 (1439, 1440) | | | | 264908 | |
| 721 | 15020180 (1441, 1442) | Novel Protein sim. GBank gjl123530 (spjP04929) HRPX_PLALO - HISTIDINE-RICH GLYCOPROTEIN PRECURSOR | | UNCLASSIFIED | 264629 | |
| 722 | 9882603 (1443, 1444) | Novel Protein sim. GBank gjl498253 (U02372) - Integrase [Vibrio cholerae] | | | 264910 | |
| 723 | 19755599 (1445, 1446) | Novel Protein sim. GBank gjl2253054 (embjCAB10705) - (Z97559) hypothetical protein Rv2114 [Mycobacterium tuberculosis] | | UNCLASSIFIED | 264691 | |
| 724 | 10126494 (1447, 1448) | Novel Protein sim. GBank gjl4083015 (AF083061) - protease P1A [Pseudomonas fluorescens] | Contains protein domain (PF00353) - Hemolysin-type calcium-binding proteins | protease | 264909 | |
| 725 | 78878679 (1449, 1450) | | | UNCLASSIFIED | 264905, 264907 | |
| 726 | 13086282 (1451, 1452) | | | UNCLASSIFIED | 264636 | |
| 727 | 13522872 (1453, 1454) | | | | 264634 | |
| 728 | 20268471 (1455, 1456) | Novel Protein sim. GBank gjl2633910 (embjCAB13411) - (Z99112) similar to hypothetical proteins [Bacillus subtilis] | | | 264587 | |
| 729 | 11293753 (1457, 1458) | | | UNCLASSIFIED | 264490 | |
| 730 | 18900373 (1459, 1460) | Novel Protein sim. GBank gjl2494860 (spjQ45291) GALE_BRELA - UDP-GLUCOSE 4-EPIMERASE (GALACTOWALDENASE) (UDP-GALACTOSE 4-EPIMERASE) | | isomerase | 264564 | |
| 731 | 80058750 (1461, 1462) | Novel Protein sim. GBank gjl1146192 (L47838) - putative [Bacillus subtilis] | | UNCLASSIFIED | 264605 | |
| 732 | 80258175 (1463, 1464) | Novel Protein sim. GBank gjl1168396 (spjP46681) AIP2_YEAST - ACTIN INTERACTING PROTEIN 2 | | struct | 264591, 264594, 264595 | |
| 733 | 20446839 (1465, 1466) | | | UNCLASSIFIED | 264604 | |
| 734 | 20435987 (1467, 1468) | Novel Protein sim. GBank gjl3184080 (embjCAA19336) - (AL023781) hypothetical protein [Schizosaccharomyces pombe] | | ubiquitin | 264604 | |

| | | | | | |
|-----|-----------------------|--|--|--------------|---|
| 735 | 11607959 (1469, 1470) | Novel Protein sim. GBank gi401382[sp]P27432[YICE_ECOLI - HYPOTHETICAL 48.9 KD PROTEIN IN GLTS-SELG INTERGENIC REGION | | | 264594 |
| 736 | 10879734 (1471, 1472) | Novel Protein sim. GBank gi400831[sp]P31135[POTH_ECOLI - PUTRESCINE TRANSPORT SYSTEM PERMEASE PROTEIN POTH | Contains protein domain (PF00528) - Binding-protein-dependent transport systems inner membrane component | transport | 264636 |
| 737 | 78945340 (1473, 1474) | | Contains protein domain (PF00815) - Regulator of G protein signaling domain | UNCLASSIFIED | 265020 |
| 738 | 17895353 (1475, 1476) | | | oxidase | 265008 |
| 739 | 78833670 (1477, 1478) | Novel Protein sim. GBank gi250666[sp]P33225[TORA_ECOLI - TRIMETHYLAMINE- N-OXIDE REDUCTASE PRECURSOR (TMAO REDUCTASE) (TRIMETHYLAMINE OXIDASE) | | | 264910 |
| 740 | 19881557 (1479, 1480) | Novel Protein sim. GBank gi3261828[emb]CAB10925] - (Z98260) mmp [Mycobacterium tuberculosis] | Contains protein domain (PF01883) - Domain of unknown function | UNCLASSIFIED | 264907, 264764, 264634, 264637 |
| 741 | 78827273 (1481, 1482) | Novel Protein sim. GBank gi3877494[emb]CAA88472.1] - (Z48583) ATP binding protein with similarity to the CDC48/PAS1/SEC18 family; cDNA EST EMBL:D65037 comes from this gene; cDNA EST EMBL:D68340 comes from this gene; cDNA EST EMBL:D65048 comes from this gene; cDNA EST EMBL:D6845... | | UNCLASSIFIED | 264689, 35696286, 264510, 264908, 18108362 |
| 742 | 82393795 (1483, 1484) | Novel Protein sim. GBank gi127420[sp]P19888[MTBA_BACAR - MODIFICATION METHYLASE BANI (CYTOSINE-SPECIFIC METHYLTRANSFERASE BANI) (M.BANI)] | Contains protein domain (PF00145) - C-5 cytosine-specific DNA methylase | | 29331822, 264910, 264762 |
| 743 | 82300051 (1485, 1486) | | | | 264488, 264259, 264508, 264905, 264906, 264907, 264908, 264909, 264510, 264511, 264512, 265008, 265009, 264910, 264591, 264596, 264759, 265010, 265011, 18108351, 264763, 264288, 264766, 264768, 264693, 18108370, 264629, 18108372, 264630, 264631, 264634, 264558, 18108385, 264482, 264564, 264567 |
| 744 | 80230421 (1487, 1488) | | | | 18108397, 264511, 264690, 264628, 264638, 264692, 264639, 264766 |
| 745 | 9841963 (1489, 1490) | Novel Protein sim. GBank gi178921[pr]J[S04846 - UDP-N- acetyl-muramoylalanine-D-glutamate-2, 6-diaminopimelate--D- alanine-D-alanine ligase (EC 6.3.2.15) precursor - Escherichia coli | | glycoprotein | 264906 |
| 746 | 11073229 (1491, 1492) | Novel Protein sim. GBank gi3386354 (AF074705) - pyocelin synthetase [Pseudomonas aeruginosa] | | synthase | 264600 |
| 747 | 94322044 (1493, 1494) | Novel Protein sim. GBank gi2887411[db]BAA24848] - (AB007878) KIAA0418 [Homo sapiens] | Contains protein domain (PF00018) - SH3 domain | oxidase | 66714117, 264905, 264509, 264906, 264907, 264908, 264909, 264511, 264910, 265011, 264681, 264288, 264766, 264687, 264768, 264769, 21906768, 35695917, 264691, 264693, 264628, 264634, 264635, 264639, 56182323, 83373044 |
| 748 | 11617923 (1495, 1496) | | | | 264690 |

| | | | | | | |
|-----|-----------------------|--|---|--------------|--|------------------|
| 749 | 20469119 (1497, 1498) | Novel Protein sim. GBank gi 1169727 sp P44948 FPG_HAEIN - FORMAMIDOPYRIMIDINE-DNA GLYCOSYLASE (FAPY- DNA GLYCOSYLASE) | | | UNCLASSIFIED | 264604 264600 |
| 750 | 20296427 (1499, 1500) | | | | | |
| 751 | 21636169 (1501, 1502) | Novel Protein sim. GBank gi 5360068 gb AAD42851.1 AF15968 - (AF15968) serine/threonine kinase PKN3 [Myxococcus xanthus] | Contains protein domain (PF00069) - Eukaryotic protein kinase domain | kinase | | 264605, 264559 |
| 752 | 82450386 (1503, 1504) | Novel Protein sim. GBank gi 1168662 sp P44426 BIOA_HAEIN - ADENOSYLMETHIONINE-8-AMINO-7-OXONONANOATE AMINOTRANSFERASE (7,8-DIAMINO-PELARGONIC ACID AMINOTRANSFERASE) (DAPA AMINOTRANSFERASE) | Contains protein domain (PF00202) - Aminotransferases class-III pyridoxal phosphate | UNCLASSIFIED | 264508, 264907, 264510, 265011, 264762, 264689, 35695855, 264638, 18108387 | |
| 753 | 80508718 (1505, 1506) | Novel Protein sim. GBank gi 2851530 sp P32399 YHGE_BACSU - HYPOTHETICAL 84.1 KD PROTEIN IN HEMY-GLTT INTERGENIC REGION (ORFB) | | UNCLASSIFIED | 264909, 264600, 264602, 264604, 264760, 264769, 264634 | |
| 754 | 95083741 (1507, 1508) | | | | | |
| 755 | 80185448 (1509, 1510) | Novel Protein sim. GBank gi 3449276 emb CAA20420 - (AL031317) putative dehydrogenase [Streptomyces coelicolor] | | UNCLASSIFIED | 264448, 264690 264769, 264689, 264638, 264639 | |
| 756 | 94631686 (1511, 1512) | | | | | |
| 757 | 79488533 (1513, 1514) | Novel Protein sim. GBank gi 4580331 emb CAB40107.1 - (AJ001206) putative glycogen debranching enzyme [Streptomyces coelicolor] | | UNCLASSIFIED | 264682, 264685 265007, 18108387, 265007, 18108387 | |
| 758 | 78963176 (1515, 1516) | | | amylase | | |
| 759 | 79475667 (1517, 1518) | Novel Protein sim. GBank gi 2911858 (AF047659) - No definition line found [Caenorhabditis elegans] | | UNCLASSIFIED | 264684, 264686 | |
| 760 | 87628888 (1519, 1520) | Novel Protein sim. GBank gi 3451312 emb CAA20449 - (AL031324) membrane altpase [Schizosaccharomyces pombe] | Contains protein domain (PF00122) - transport E1-E2 ATPase | | 29331822, 264908, 52644045, 56182435, 60170831, 21906754, 265017, 265019, 264681, 264687, 264688, 21906766, 21906768, 265020, 265021, 265022, 264635, 22279000 | |
| 761 | 79877966 (1521, 1522) | Novel Protein sim. GBank gi 3327158 db BAA31647 - (AB014572) KIAA0672 protein [Homo sapiens] | | UNCLASSIFIED | 264766 | |
| 762 | 80023563 (1523, 1524) | | | UNCLASSIFIED | 264907, 264593, 265020 | |
| 763 | 20294813 (1525, 1526) | Novel Protein sim. GBank gi 4981266 gb AAC35822.1 AE00174 - (AE001744) lipopolysaccharide core biosynthesis protein KdIB [Thermotoga maritima] | | | 264600 | |
| 764 | 39515024 (1527, 1528) | | | | | 264603 |

| | | | | | |
|-----|-----------------------|---|---|--------------|--|
| 765 | 80025347 (1529, 1530) | Novel Protein sim. GBank gij3845093 (AE001371) - erythrocyte membrane protein PIEMP3 [Plasmodium falciparum] | | struct | 264805, 264906, 264594, 264686, 33657023 |
| 766 | 82417404 (1531, 1532) | Novel Protein sim. GBank gij541121 [pir]S40827 - hypothetical protein o300 - Escherichia coli | | UNCLASSIFIED | 264605, 264762, 18108374 |
| 767 | 10296742 (1533, 1534) | Novel Protein sim. GBank gij2982501 [emb]CAA061641 - hypothetical protein - Trypanosoma cruzi (fragment) | | UNCLASSIFIED | 264907 |
| 768 | 79416080 (1535, 1536) | Novel Protein sim. GBank gij2982501 [emb]CAA061641 - hypothetical protein - Trypanosoma cruzi (fragment) | | esterase | 264592, 264595 |
| 769 | 80086554 (1537, 1538) | Novel Protein sim. GBank gij2982501 [emb]CAA061641 - hypothetical protein - Trypanosoma cruzi (fragment) | | UNCLASSIFIED | 55810764, 264559 |
| 770 | 80417847 (1539, 1540) | Novel Protein sim. GBank gij2982501 [emb]CAA061641 - hypothetical protein - Trypanosoma cruzi (fragment) | | UNCLASSIFIED | 264905, 264907, 264828, 264909, 265010, 264766, 264628, 264629, 264634, 264636, 264555 |
| 771 | 95329509 (1541, 1542) | Novel Protein sim. GBank gij4769004 [gb]AAD29715.1 [AF140598] ring-box protein 1 [Homo sapiens] | Contains protein domain (PF000097) - Zinc finger, C3HC4 type (RING finger) | UNCLASSIFIED | 56182575, 35696286, 264259, 66714117, 264107, 66712502, 56182435, 264112, 55812038, 87168559, 264288, 21906766, 33657023, 65274620, 65274791, 18108381 |
| 772 | 76971362 (1543, 1544) | Novel Protein sim. GBank gij5531324 [emb]CAB51045.1 - putative alkane 1-monoxygenase [Pseudomonas fluorescens] | | UNCLASSIFIED | 264910 |
| 773 | 76945363 (1545, 1546) | Novel Protein sim. GBank gij5531324 [emb]CAB51045.1 - putative alkane 1-monoxygenase [Pseudomonas fluorescens] | | UNCLASSIFIED | 265020 |
| 774 | 79856129 (1547, 1548) | Novel Protein sim. GBank gij5531324 [emb]CAB51045.1 - putative alkane 1-monoxygenase [Pseudomonas fluorescens] | | UNCLASSIFIED | 264909 |
| 775 | 20620141 (1549, 1550) | Novel Protein sim. GBank gij4505461 [ref]NP_003624.1 [pNRPB - nuclear restricted protein, BTB domain-like (brain)] | Contains protein domain (PF01006) - Hepatitis C virus non-structural protein NS4a | UNCLASSIFIED | 264555 |
| 776 | 76942693 (1551, 1552) | Novel Protein sim. GBank gij4505461 [ref]NP_003624.1 [pNRPB - nuclear restricted protein, BTB domain-like (brain)] | Contains protein domain (PF01344) - Kelch motif | protease | 265019 |
| 777 | 79960378 (1553, 1554) | Novel Protein sim. GBank gij4505461 [ref]NP_003624.1 [pNRPB - nuclear restricted protein, BTB domain-like (brain)] | | UNCLASSIFIED | 21906754, 265020, 60170615, 264691 |
| 778 | 20691310 (1555, 1556) | Novel Protein sim. GBank gij1144520 (U34956) - phosphoribosylformylglycinamide synthase [Mycobacterium tuberculosis] | | UNCLASSIFIED | 264511 |
| 779 | 80054024 (1557, 1558) | Novel Protein sim. GBank gij1144520 (U34956) - phosphoribosylformylglycinamide synthase [Mycobacterium tuberculosis] | | synthase | 264603 |
| 780 | 95288987 (1559, 1560) | Novel Protein sim. GBank gij1144520 (U34956) - phosphoribosylformylglycinamide synthase [Mycobacterium tuberculosis] | | | 264907, 264600, 264601, 264602, 264603, 264604, 264605, 264486 |
| 781 | 80250049 (1561, 1562) | Novel Protein sim. GBank gij4155447 (AE001517) - proline/betaine transporter [Helicobacter pylori J99] | | UNCLASSIFIED | 264905, 264907, 265010, 264600, 264601, 18108362, 18108374, 264556 |
| 782 | 8758529 (1563, 1564) | Novel Protein sim. GBank gij4155447 (AE001517) - proline/betaine transporter [Helicobacter pylori J99] | | UNCLASSIFIED | 264605 |
| 783 | 16410791 (1565, 1566) | Novel Protein sim. GBank gij4155447 (AE001517) - proline/betaine transporter [Helicobacter pylori J99] | | UNCLASSIFIED | 265020 |
| 784 | 80051197 (1567, 1568) | Novel Protein sim. GBank gij4155447 (AE001517) - proline/betaine transporter [Helicobacter pylori J99] | | UNCLASSIFIED | 264635, 33657023, 29331828, 265017, 264565, 264566 |
| 785 | 56073541 (1569, 1570) | Novel Protein sim. GBank gij3451335 (AC005525) - F22162_1 [Homo sapiens] | Contains protein domain (PF00047) - Immunoglobulin domain | struct | 35696052, 264604 |
| 786 | 20438842 (1571, 1572) | Novel Protein sim. GBank gij136748 [sp]P10905 [UGPA_ECOLI - SN-GLYCEROL-3-PHOSPHATE TRANSPORT SYSTEM PERMEASE PROTEIN UGPA] | | transport | 264603 |
| 787 | 80258364 (1573, 1574) | Novel Protein sim. GBank gij136748 [sp]P10905 [UGPA_ECOLI - SN-GLYCEROL-3-PHOSPHATE TRANSPORT SYSTEM PERMEASE PROTEIN UGPA] | | UNCLASSIFIED | 264593 |

| | | | | | |
|-----|-----------------------|---|--|------------------|--|
| 788 | 80507844 (1575, 1576) | Novel Protein sim. GBank gij2746079 (AF015310) - BTH1 [Brassica napus] | | synthase | 264908, 264602, 264603, 264769, 264638 |
| 789 | 17294715 (1577, 1578) | Novel Protein sim. GBank gij2351849 (U933357) - 40 kDa heat shock chaperone protein [Halobacterium salinarum] | | UNCLASSIFIED | 265007 |
| 790 | 86284406 (1579, 1580) | Novel Protein sim. GBank gij5706378 [dbj]BAA83099.1 - (AB026118) MALT1 [Homo sapiens] | Contains protein domain (PF00047) - Immunoglobulin domain | glycoprotein | 22279899, 28331824, 264828, 87188559, 265018, 21906765, 21906767, 21906768, 21906769, 265020, 264692, 22279000, 264563 |
| 791 | 94851627 (1581, 1582) | Novel Protein sim. GBank gij5689948 [emb]CAB51985.1 - (AL109863) putative isoleucyl-tRNA synthetase [Streptomyces coelicolor A3(2)] | | | 264601, 264605, 264636 |
| 792 | 80058786 (1583, 1584) | Novel Protein sim. GBank gij393194 (L023375) - S-antigen [Plasmodium falciparum] | | struct | 265021, 264631, 264635, 264556 |
| 793 | 79638730 (1585, 1586) | Novel Protein sim. GBank gij1345408 [dbj]BAA05046 - (D26046) AT motif-binding factor [Mus musculus] | Contains protein domain (PF00046) - Homeobox domain | homeobox | 264693 |
| 794 | 81839294 (1587, 1588) | Novel Protein sim. GBank gij105884 [pir]S24023 - dopamine receptor D4 - human (fragment) | | UNCLASSIFIED | 264603, 264604, 264910, 264762, 264906, 264639, 264909, 264757 |
| 795 | 80074988 (1589, 1590) | Novel Protein sim. GBank gij1877334 [emb]CAB07082 - (Z92771) birA [Mycobacterium tuberculosis] | | carboxylase | 264488, 35696052, 264905, 264907, 265010, 35696423, 264636 |
| 796 | 86569451 (1591, 1592) | | | | 60432229, 55811150, 264630, 264637, 264565 |
| 797 | 87771781 (1593, 1594) | Novel Protein sim. GBank gij2995447 [emb]CAA71519 - (Y10495) CDV-1R protein [Mus musculus] | | struct | 22278998, 264093, 264094, 66714117, 21906767, 21906769, 265020, 265022 |
| 798 | 79865209 (1595, 1596) | | | transcriptfactor | 264687, 264768, 264693 |
| 799 | 79857816 (1597, 1598) | Novel Protein sim. GBank gij4467250 [emb]CAB37575 - (AL035569) probable Glu-tRNA Gln amidotransferase subunit [Streptomyces coelicolor] | | hydrolase | 264909, 264910, 264636, 264638 |
| 800 | 79970189 (1599, 1600) | | | | |
| 801 | 80499399 (1601, 1602) | Novel Protein sim. GBank gij2791517 [emb]CAA16054 - (AL021246) hypothetical protein Rv2477c [Mycobacterium tuberculosis] | | UNCLASSIFIED | 264488 |
| 802 | 79834598 (1603, 1604) | Novel Protein sim. GBank gij4887211 [gb]AAD32237.1 [AF14744 - (AF147449) penicillin binding protein 1B [Pseudomonas aeruginosa] | | transport | 264508, 264511, 265006, 265009, 264769, 264567, 264486 |
| 803 | 20467520 (1605, 1606) | | | struct | 264605 |
| 804 | 10174239 (1607, 1608) | Novel Protein sim. GBank gij1176152 [sp]P44507 [YHAD_HAEIN - HYPOTHETICAL PROTEIN H0091] | | kinase | 264510 |
| 805 | 79599993 (1609, 1610) | | | | |
| 806 | 80484113 (1611, 1612) | Novel Protein sim. GBank gij2764612 [emb]CAA04683 - (AJ001330) ornithine transcarbamoylase [Lactobacillus sakei] | Contains protein domain (PF00185) - Aspartate/ornithine carbamoyltransferase | - transferase | 264508, 264769 |
| 807 | 80381812 (1613, 1614) | Novel Protein sim. GBank gij2833311 [sp]Q21828 [YNF0 CAEEL - HYPOTHETICAL 18.9 KD PROTEIN R07E5.13 IN CHROMOSOME III] | | | 264764 |

| | | | | | |
|-----|-----------------------|--|--|--------------|--|
| 808 | 35106817 (1615, 1616) | Novel Protein sim. GBank gij3913092[sp]Q46170[ARCD_CLOPE - ARGININE/ORNITHINE ANTIporter] | | transport | 264909, 264602, 21906764, 18108374 |
| 809 | 81454254 (1617, 1618) | Novel Protein sim. GBank gij3913016[sp]P74309[ALF1_SNNY3 - FRUCTOSE- BISPHOSPHATE ALDOLASE CLASS I (FBP ALDOLASE)] | Contains protein domain (PF00274) - Fructose-bisphosphate aldolase class-I | UNCLASSIFIED | 264508, 264906, 264909, 265007, 264910, 264758, 264600, 264602, 264603, 264605, 264687, 264769, 264689, 264636, 264486 |
| 810 | 80192761 (1618, 1620) | Novel Protein sim. GBank gij401472[sp]P30863[YAFB_ECOLI - HYPOTHETICAL OXIDOREDUCTASE IN ASPU-MLTD INTERGENIC REGION] | Contains protein domain (PF00248) - Aldoketo reductase family | reductase | 284369 |
| 811 | 80079280 (1621, 1622) | | | UNCLASSIFIED | 264558 |
| 812 | 10297654 (1623, 1624) | | | UNCLASSIFIED | 264692 |
| 813 | 79612280 (1625, 1626) | | | UNCLASSIFIED | 264906 |
| 814 | 80473427 (1627, 1628) | Novel Protein sim. GBank gij146168 (J01617) - glutaminyl- [RNA synthetase [Escherichia coli]] | | synthase | 264905, 264602, 264605, 264682, 264687, 264769, 264636 |
| 815 | 95419513 (1629, 1630) | Novel Protein sim. GBank gij14589652[dbj]BAA76848.1] - (AB023221) KIAA1004 protein [Homo sapiens] | | UNCLASSIFIED | 264488, 22278998, 22278999, 29331822, 29331824, 29331825, 29331827, 29331828, 29146489, 264905, 264908, 265007, 33657402, 60433356, 60433438, 264758, 265011, 265017, 265018, 265019, 264369, 264288, 264685, 21906765, 21906767, 265020, 265021, 264892, 65274620, 33657109, 264629, 18108376, 264635, 264638, 60170394, 58182323, 264564 264600 |
| 816 | 18881910 (1631, 1632) | | | UNCLASSIFIED | 264595 |
| 817 | 95293316 (1633, 1634) | Novel Protein sim. GBank gij1781144[emb]CAB06254] - (Z83866) hypothetical protein Rv3069 [Mycobacterium tuberculosis] | | | |
| 818 | 90938190 (1635, 1636) | Novel Protein sim. GBank gij1477468 (U35244) - vacuolar protein sorting homolog r-vps33a [Rattus norvegicus] | | | 65274572, 22278999, 60424269, 35696052, 55812038, 21906768, 55811957, 35695917, 33657023, 18108370, 18108374, 55810764, 35696423, 55811578, 264638 |
| 819 | 80254977 (1637, 1638) | Novel Protein sim. GBank gij1001352[dbj]BAA10839] - (D64008) ABC transporter [Synechocystis sp.] | | transport | 264585 |
| 820 | 80059688 (1639, 1640) | Novel Protein sim. GBank gij586814[sp]P37484[YVT_BACSU - HYPOTHETICAL 74.3 KD PROTEIN IN RPLI-COTF INTERGENIC REGION] | | UNCLASSIFIED | 264600, 264602, 264604 |
| 821 | 79762590 (1641, 1642) | | | UNCLASSIFIED | 264910 |
| 822 | 80215310 (1643, 1644) | | | UNCLASSIFIED | 264510, 264594, 264637 |
| 823 | 94992299 (1645, 1646) | Novel Protein sim. GBank gij3878400[emb]CAA95828] - (Z71284) predicted using GeneFinder; Weak similarity to Mouse T-complex-associated-testes-expressed-1 protein (PIR Acc. No. A45841); cDNA EST EMBL:D32742 comes from this gene; cDNA EST EMBL:D33617 comes from this gene; cDNA EST.... | | struct | 264509, 264687, 264691 |
| 824 | 80411171 (1647, 1648) | Novel Protein sim. GBank gij1370076[emb]CAA66887] - (X98235) type I [Drosophila melanogaster] | Contains protein domain (PF01429) - Methyl-CpG binding domain | | 264910, 264763, 264769, 264693 |

| | | | | | | | |
|-----|-----------------------|--|--|--|--------------|--|--|
| 825 | 20638600 (1648, 1650) | Novel Protein sim. GBank gil3025132[sp]P77391[YEAG_ECOLI - HYPOTHETICAL 74.5 KD PROTEIN IN GAPA-RND INTERGENIC REGION (AL021646) hypothetical protein Rv3202c [Mycobacterium tuberculosis] | | | UNCLASSIFIED | 264592 | |
| 826 | 11075047 (1651, 1652) | Novel Protein sim. GBank gil324228[lemb]CAA16669] - (AL021646) hypothetical protein Rv3202c [Mycobacterium tuberculosis] | | | | 264605 | |
| 827 | 60054207 (1653, 1654) | Novel Protein sim. GBank gil3417424[lemb]CAA20312] - (AL031261) putative transport protein [Schizosaccharomyces pombe] | | | | 264603 | |
| 828 | 95106322 (1655, 1656) | Novel Protein sim. GBank gil4336692[gb]AAD17897] - (AF101361) Abnormal X segregation [Drosophila melanogaster] | | | UNCLASSIFIED | 52645080, 264508, 264905, 264509, 264906, 264807, 264908, 264909, 264512, 264910, 264591, 264758, 264800, 264768, 264768, 21908768, 35695917, 284691, 264628, 264630, 264631, 264632, 264634, 264635, 264636, 264637, 264558, 264639, 83373044, 18108385, 264563, 264566, 264486 | |
| 829 | 81742215 (1657, 1658) | Novel Protein sim. GBank gil3820539 (AF080002) - UDP-N- acetylmutaric tripeptide synthetase MurC [Heliobacillus mobilis] | | | UNCLASSIFIED | 264758, 264634 | |
| 830 | 20396091 (1659, 1660) | Novel Protein sim. GBank gil1870004[lemb]CAB06855] - (Z82539) hypothetical protein Rv1024 [Mycobacterium tuberculosis] | | | UNCLASSIFIED | 66714117, 264910, 264639 | |
| 831 | 87112435 (1661, 1662) | Novel Protein sim. GBank gil2500056[sp]Q46267[PFLA_CLOPA - PYRUVATE FORMATE-LYASE ACTIVATING ENZYME | | | UNCLASSIFIED | 264602 | |
| 832 | 19536322 (1663, 1664) | Novel Protein sim. GBank gil2497531[sp]Q46078[KPYK_CORGL - PYRUVATE KINASE (PK) | | Contains protein domain (PF00224) - Pyruvate kinase | UNCLASSIFIED | 264600, 264602, 264769, 264689, 264636 | |
| 833 | 20726654 (1665, 1666) | Novel Protein sim. GBank gil2496533[sp]Q50598[Y0D8_MYCTU - HYPOTHETICAL 69.9 KD PROTEIN CY1A11.08 | | | | 264768, 263994, 21908767, 264910, 264632, 264635, 264259, 264639, 264693, 83373044, 264758, 35696052, 22279002, 264508, 264905, 264906, 264448, 263972, 264908, 264909 | |
| 834 | 21428762 (1667, 1668) | Novel Protein sim. GBank gil4557753[re]NP_000372.1[pmid1 - midline 1 protein finger] | | | transport | 35695917, 264557 | |
| 835 | 94140482 (1669, 1670) | Novel Protein sim. GBank gil728867[sp]P40602[APG_ARATH - ANTER-SPECIFIC PROLINE-RICH PROTEIN APG PRECURSOR | | | UNCLASSIFIED | 264595 | |
| 836 | 56126552 (1671, 1672) | Novel Protein sim. GBank gil2496533[sp]Q50598[Y0D8_MYCTU - HYPOTHETICAL 69.9 KD PROTEIN CY1A11.08 | | | UNCLASSIFIED | 264687 | |
| 837 | 79450450 (1673, 1674) | Novel Protein sim. GBank gil4557753[re]NP_000372.1[pmid1 - midline 1 protein finger] | | | UNCLASSIFIED | 264806 | |
| 838 | 79184203 (1675, 1676) | Novel Protein sim. GBank gil4557753[re]NP_000372.1[pmid1 - midline 1 protein finger] | | | | 264762, 264556 | |
| 839 | 78641125 (1677, 1678) | Novel Protein sim. GBank gil4557753[re]NP_000372.1[pmid1 - midline 1 protein finger] | | | | | |
| 840 | 60059851 (1679, 1680) | Novel Protein sim. GBank gil4557753[re]NP_000372.1[pmid1 - midline 1 protein finger] | | | | | |

| | | | | | |
|-----|-----------------------|---|---|------------------|---|
| 841 | 80376318 (1681, 1682) | Novel Protein sim. GBank glt139805[sp]P08045[XFIN_XENLA - XFIN PROTEIN (D88733)] membrane glycoprotein [Equine herpesvirus 1] | Contains protein domain (PF00096) Zinc finger, C2H2 type | transcriptfactor | 264764 |
| 842 | 80078724 (1683, 1684) | Novel Protein sim. GBank glt2114321[dbj]BAA200371 - (D88733) membrane glycoprotein [Equine herpesvirus 1] | Contains protein domain (PF00569) Zinc finger present in dystrophin, CBP/p300 | UNCLASSIFIED | 264905, 264908, 265008, 265009, 18108374, 56182323, 264558 |
| 843 | 87002847 (1685, 1686) | Novel Protein sim. GBank glt3882325[dbj]BAA34522.1] - (AB018345) KIAA0802 protein [Homo sapiens] | Contains protein domain (PF00170) bZIP transcription factor | strut | 264091, 29331825, 264906, 264768, 264563 |
| 844 | 17841439 (1687, 1688) | Novel Protein sim. GBank glt224721[dbj]BAA20844] - (AB002388) KIAA0390 [Homo sapiens] | Contains protein domain (PF00096) Zinc finger, C2H2 type | transcriptfactor | 265011 |
| 845 | 18346844 (1688, 1690) | | | | 264629 |
| 846 | 79863441 (1691, 1692) | Novel Protein sim. GBank glt625679[pir]A36929 - virulence regulatory protein VsrB - Pseudomonas solanacearum | | Kinase | 264907 |
| 847 | 78695348 (1693, 1694) | | | | 264909 |
| 848 | 78489365 (1695, 1696) | | | UNCLASSIFIED | 265020 |
| 849 | 78756367 (1697, 1698) | | | UNCLASSIFIED | 264568 |
| 850 | 78817849 (1699, 1700) | Novel Protein sim. GBank glt3183245[sp]P78061[JC_ECOLI - PUTATIVE GLUTAMINE SYNTHETASE (GLUTAMATE-AMMONIA LIGASE)] | Contains protein domain (PF00120) Glutamine synthetase | UNCLASSIFIED | 264909 |
| 851 | 95320333 (1701, 1702) | Novel Protein sim. GBank glt5454130[re]NP_008280.1p[TLN] - talin | Contains protein domain (PF01608) ILWEQ domain | | 264488, 52644507, 264489, 18108398, 65274572, 56182575, 22278994, 22278995, 22278996, 35696286, 22278997, 22278998, 22278999, 20281171, 264490, 264259, 52645080, 29331822, 29331824, 66714117, 29331825, 60432289, 29331826, 29331827, 35696052, 29331828, 29146498, 29146499, 264107, 264905, 264906, 264907, 264908, 52644045, 56182435, 265006, 265007, 265008, 265009, 264910, 60432229, 60431735, 60433356, 33657402, 60433438, 264595, 264758, 264759, 21906754, 33109954, 52644296, 265010, 265011, 87168559, 265017, 265018, 265019, 264760, 264761, 264762, 264681, 18108351, 264763, 264448, 264682, 264784, 264683, 18108354, 264288, 264369, 264685, 264766, 264687, 264768, 264769, 21906765, 21906766, 21906767, 21906768, 29148627, 21906769, 29148629, 55811857, 35695917, 265020, 265021, 265022, 60170615, 52644150, 264691, 264692, 33657023, 264693, 263966, 33657109, 27486261, 27486262, 27486264, 27486265, 35695763, 60431802, 18108370, 20281089, 264629, 18108374, 18108376, 55811576, 35696423, 35695855, 264634, 264635, 264636, 264555, 60431850, 264556, 264691 |
| 852 | 10147366 (1703, 1704) | | | | |

| | | | | | |
|-----|-----------------------|--|--|--------------|--|
| 853 | 13032587 (1705, 1706) | Novel Protein sim. GBank gij3402836[embj][CAA76082] - (Y16136) 2-enoate reductase [Moorella thermoacelica] | | reductase | 264636 |
| 854 | 80052438 (1707, 1708) | | | | 264566 |
| 855 | 79641130 (1709, 1710) | | | | 264692 |
| 856 | 11594238 (1711, 1712) | | | UNCLASSIFIED | 264591 |
| 857 | 79210165 (1713, 1714) | | | UNCLASSIFIED | 264630, 264634 |
| 858 | 80248910 (1715, 1716) | | | | 265008, 265009, 264601, 264602, 264603, 18108351 |
| 859 | 20296634 (1717, 1718) | | | | 264559 |
| 860 | 80041749 (1718, 1720) | | | UNCLASSIFIED | 264489 |
| 861 | 85857045 (1721, 1722) | | | UNCLASSIFIED | 33857023, 264630 |
| 862 | 80079467 (1723, 1724) | | | | 264600 |
| 863 | 80578931 (1725, 1726) | Novel Protein sim. GBank gij2246532 (U93872) - ORF 73, contains large complex repeat CR 73 [Kaposi's sarcoma-associated herpesvirus] | | UNCLASSIFIED | 264488, 18108398, 35896286, 264259, 18108351, 264288, 265021 |
| 864 | 94839904 (1727, 1728) | | | UNCLASSIFIED | 264259, 264112, 263974 |
| 865 | 80045310 (1729, 1730) | Novel Protein sim. GBank gij5689884[embj][CAB52047.1] - (AL109732) hypothetical protein [Streptomyces coelicolor A3(2)] | Contains protein domain (PF01479) - S4 domain | | 264635, 264600, 264636, 264591, 264602, 264693 |
| 866 | 80162031 (1731, 1732) | Novel Protein sim. GBank gij455787[ref][NP_000341.1]pABCR - ATP binding cassette transporter | | transport | 264288, 264557, 264558 |
| 867 | 80062402 (1733, 1734) | | | | 264605 |
| 868 | 10075364 (1735, 1736) | | | UNCLASSIFIED | 264909 |
| 869 | 80062406 (1737, 1738) | | | | 264605, 264687, 18108374 |
| 870 | 80248651 (1739, 1740) | Novel Protein sim. GBank gij628660[pir][S37755 - Adenyl- transferase - Escherichia coli | | transferase | 264601, 264636 |
| 871 | 20378285 (1741, 1742) | Novel Protein sim. GBank gij1708180[sp]Q10602[HEMK_MYCTU - HEMK PROTEIN HOMOLOG | | UNCLASSIFIED | 264603 |
| 872 | 95197114 (1743, 1744) | Novel Protein sim. GBank gij1545959[embj][CAA67763] - (X99384) paladin [Mus musculus] | | UNCLASSIFIED | 35896286, 22278998, 264259, 29331822, 29331824, 29331825, 60432289, 29331826, 29331827, 29331828, 35896052, 264509, 264905, 264906, 264907, 264908, 264909, 264510, 265006, 264511, 264512, 265007, 265008, 265009, 264910, 264591, 60433356, 264596, 52646317, 87168474, 265010, 264602, 264603, 265017, 265018, 264605, 18108351, 264764, 264766, 264768, 52644229, 264769, 21906765, 265021, 264534, 264691, 52645129, 264628, 264629, 35898423, 65274781, 264631, 264632, 264635, 264636, 264556, 264637, 264638, 264639, 60432113, 22279000, 22279002, 264594 |
| 873 | 20189728 (1745, 1746) | Novel Protein sim. GBank gij4156104 (AE001569) - putative Outer membrane protein [Helicobacter pylori J99] | Contains protein domain (PF00096) - Zinc finger, C2H2 type | UNCLASSIFIED | 264595 |

| | | | | | |
|-----|-----------------------|---|--|--------------|--|
| 874 | 80077692 (1747, 1748) | Novel Protein sim. GBank gii134319 sp P07819 SCRB_BACSU - SUCROSE-6- PHOSPHATE HYDROLASE (SUCRASE) (INVERTASE) Novel Protein sim. GBank gii481000 p j S37594 - mucin - human (fragment) | | UNCLASSIFIED | 264600 |
| 875 | 86608446 (1748, 1750) | Novel Protein sim. GBank gii481000 p j S37594 - mucin - human (fragment) | | | 264259, 264448, 264288, 264557, 87168518 |
| 876 | 86465157 (1751, 1752) | Novel Protein sim. GBank gii3128283 (AF010496) - iron(iii) dicitrate transport ATP-binding protein [Rhodobacter capsulatus] | Contains protein domain (PF000005) - transport ABC transporter | | 264907, 264601, 264602, 264605, 265020, 60431602 |
| 877 | 87802548 (1753, 1754) | Novel Protein sim. GBank gii731074 sp P40349 URB1_USTMA - SIDEROPHORE BIOSYNTHESIS REGULATORY PROTEIN URBS1 | Contains protein domain (PF00320) - transcriptfactor GATA zinc finger | | 22278998, 264908, 264369 |
| 878 | 80187289 (1755, 1756) | Novel Protein sim. GBank gii1351614 sp Q09853 YAEED_SCHPO - HYPOTHETICAL 181.5 KD PROTEIN C23D3.13C IN CHROMOSOME 1 | ATPase_associated | | 264369, 264555 |
| 879 | 94328962 (1757, 1758) | Novel Protein sim. GBank gii3875304 emb CAA98434 - (Z74030) predicted using GeneFinder; cDNA EST EMBL: C07609 comes from this gene; cDNA EST EMBL: C09023 comes from this gene; cDNA EST yk505e9.3 comes from this gene; cDNA EST yk489h9.3 comes from this gene; cDNA EST yk489h9.5 com... | | | 56182575, 29331824, 264508, 264906, 265018, 18108351, 264448, 264683, 21908768, 21908768, 60170815, 33657023, 65274620, 33657109, 18108374, 35695855, 264563 |
| 880 | 8491135 (1759, 1760) | Novel Protein sim. GBank gii137120 sp P11214 UR0T_MOUSE - TISSUE PLASMINOGEN ACTIVATOR PRECURSOR (TPA) (T- PLASMINOGEN ACTIVATOR) | Contains protein domain (PF00051) - cathepsin Kringle domain | | 264508 |
| 881 | 11290122 (1761, 1762) | | | UNCLASSIFIED | 264508 |
| 882 | 11077011 (1763, 1764) | Novel Protein sim. GBank gii2632098 emb CAA75667 - (Y15513) Prodos protein [Drosophila melanogaster] | | UNCLASSIFIED | 264558 |
| 883 | 79582968 (1765, 1766) | Novel Protein sim. GBank gii1155068 emb CAA64425 - (X94976) cell wall-plasma membrane linker protein [Brassica napus] | | UNCLASSIFIED | 264688 |
| 884 | 13517921 (1767, 1768) | Novel Protein sim. GBank gii2078027 emb CAB08467 - (Z95208) hypothetical protein Rv2372c [Mycobacterium tuberculosis] | | UNCLASSIFIED | 264636 |
| 885 | 80052457 (1769, 1770) | | | UNCLASSIFIED | 264605, 18108362 |
| 886 | 11685136 (1771, 1772) | Novel Protein sim. GBank gii2695834 emb CAA15904 - (AL021006) sucA [Mycobacterium tuberculosis] | | | 264690 |
| 887 | 94315307 (1773, 1774) | Novel Protein sim. GBank gii5689395 dbj BAA82981.1 - (AB028952) KIAA1029 protein [Homo sapiens] | dehydrogenase | | 35696052, 264906, 264600, 264603, 35695917, 35695855, 264636 |
| 888 | 10083399 (1775, 1776) | Novel Protein sim. GBank gii1881338 dbj BAA19365 - (AB001488) PROBABLE INTEGRAL MEMBRANE PROTEIN, SIMILAR TO CHLORAMPHENICOL RESISTANCE PROTEIN OF STREPTOMYCES VENEZUELAE [Bacillus subtilis] | UNCLASSIFIED | | 264908 |
| 889 | 20385917 (1777, 1778) | Novel Protein sim. GBank gii854065 emb CAA58337 - Novel Protein sim. GBank gii854065 emb CAA58337 - (X83413) U88 [Human herpesvirus 6] | | | 264603 |
| 890 | 19904337 (1779, 1780) | | | | 264629 |

| | | | | | |
|-----|-----------------------|---|--|------------------|--|
| 891 | 13516879 (1781, 1782) | Novel Protein sim. GBank gi 4959398 gb AAD34331.1 AF11248 - (AF11248) RAD54B protein [Homo sapiens] | | UNCLASSIFIED | 264636 |
| 892 | 87634157 (1783, 1784) | Novel Protein sim. GBank gi 545526 bbs 143833 - LBP- 1b=transcription factor binding to initiation site of HIV-1 (alternatively spliced) [human, Namahwa cells, Peptide, 541 aa] | | transcriptfactor | 22278986, 22278999, 29331828, 35696052, 264908, 264909, 265009, 265011, 264602, 265018, 264766, 21908765, 21908768, 21908769, 265020, 265021, 56526486 264689, 263967 |
| 893 | 79168037 (1785, 1786) | Novel Protein sim. GBank gi 2829688 sp P80508 CYSK_MAIZE - CYSTEINE SYNTHASE (O-ACETYL-SERINE SULFHYDRYLASE) (O- ACETYL-SERINE (THIOL)-LYASE) (CSASE) | | synthase | |
| 894 | 11102240 (1787, 1788) | | | | 263978 |
| 895 | 80239868 (1789, 1790) | | | UNCLASSIFIED | 264508, 264600, 264555, 264559 |
| 896 | 79747803 (1791, 1792) | | | | 264632 |
| 897 | 94991923 (1793, 1794) | | | UNCLASSIFIED | 264686, 29331828, 264511 |
| 898 | 87895109 (1795, 1796) | | Contains protein domain (PF00047) - Immunoglobulin domain | | 56182575, 60432289, 56182435, 60432229, 55811957, 22279000, 264486 264601 |
| 899 | 11100463 (1797, 1798) | Novel Protein sim. GBank gi 1750127 (U66480) - YncC [Bacillus subtilis] | | transport | 264769, 264691, 264563 |
| 900 | 80499768 (1799, 1800) | Novel Protein sim. GBank gi 3122879 sp O07438 SYA_MYCTU - ALANYL-TRNA SYNTHETASE (ALANINE-TRNA LIGASE) (ALARS) | Contains protein domain (PF01411) - tRNA synthetases class II (A) | synthase | 264907, 264602, 264605, 264769, 35695917, 18108376, 264563 |
| 902 | 80503301 (1803, 1804) | Novel Protein sim. GBank gi 335570 emb CAA20001 - (AL031124) 3-Isopropylmalate dehydratase large subunit [Streptomyces coelicolor] | Contains protein domain (PF00330) - Aconitase family (aconitase hydratase) | isomerase | 264909, 265008, 264602, 264604, 264769, 264689, 264693 |
| 903 | 82060208 (1805, 1806) | Novel Protein sim. GBank gi 2960120 emb CAA18018.1 - (AL022121) glpK [Mycobacterium tuberculosis] | Contains protein domain (PF00370) - FGG family of carbohydrate kinases | kinase | 35996052, 264905, 264510, 264511, 264512, 264605, 264760, 18108351, 264762, 264687, 264768, 264769, 264688, 21908764, 35695917, 27486262, 35695955, 264634, 264638, 264486 264604 |
| 904 | 20451078 (1807, 1808) | Novel Protein sim. GBank gi 728887 sp P40506 ARGI_COCIM - ARGINASE | Contains protein domain (PF00491) - Arginase family | hydrolase | |
| 905 | 9398483 (1809, 1810) | Novel Protein sim. GBank gi 4567200 gb AAD23616.1 AC00718 - (AC00718) hypothetical protein [Arabidopsis thaliana] | | UNCLASSIFIED | 264809 |
| 906 | 80052628 (1811, 1812) | Novel Protein sim. GBank gi 3738200 emb CAA21292 - (AL031855) putative vacuolar membrane protein [Schizosaccharomyces pombe] | | | 264595, 264605 |
| 907 | 87913201 (1813, 1814) | | | UNCLASSIFIED | 60432289, 264601, 264690 |
| 908 | 11754482 (1815, 1816) | | | UNCLASSIFIED | 264638 |
| 909 | 20727907 (1817, 1818) | Novel Protein sim. GBank gi 3668940 dbj BAA34296 - (AB015054) Alg2 [Rhizomucor pusillus] | | UNCLASSIFIED | 264602 |
| 910 | 16776206 (1819, 1820) | Novel Protein sim. GBank gi 4569726 dbj BAA76883.1 - (AB003137) DnaJ homolog protein [Salix glauca] | Contains protein domain (PF00684) - DnaJ central domain (4 repeats) | eph | 265009 |

| | | | | | |
|-----|-----------------------|--|---|---------------|--|
| 911 | 87454340 (1821, 1822) | Novel Protein sim. GBank gi 548774 sp P35685 RL7A_ORYSA - 60S RIBOSOMAL PROTEIN L7A | | ribosomalprot | 265010, 264604, 60432113 |
| 912 | 20448863 (1823, 1824) | Novel Protein sim. GBank gi 2314008 gb AAD07921.1 - (AE000597) CDP-diglyceride hydrolase (cdh) [Helicobacter pylori 26695] | | hydrolase | 284559 |
| 913 | 20469357 (1825, 1826) | Novel Protein sim. GBank gi 417657 sp Q03604 RIR1_CAEEL - PROBABLE RIBONUCLEOSIDE-DIPHOSPHATE REDUCTASE LARGE CHAIN (RIBONUCLEOTIDE REDUCTASE) | Contains protein domain (PF00317) - Ribonucleotide reductase | UNCLASSIFIED | 264604 |
| 914 | 79183351 (1827, 1828) | Novel Protein sim. GBank gi 5689571 dbj BAA83069.1 - (AB029040) KIAA1117 protein [Homo sapiens] | | reductase | 264636 |
| 915 | 87606703 (1829, 1830) | Novel Protein sim. GBank gi 186110 emb CAA71790 - (Y10831) putative integrase [Ralstonia eutropha] | | UNCLASSIFIED | 18108398, 22278996, 66714117, 264906, 284591, 21908768, 265020, 55811576, 264638 |
| 916 | 79444091 (1831, 1832) | Novel Protein sim. GBank gi 1655699 emb CAA69032 - (Y07752) pterophorin-S [Volvox carter] | | UNCLASSIFIED | 264595 |
| 917 | 20195985 (1833, 1834) | Novel Protein sim. GBank gi 5689968 emb CAB52005.1 - (AL109663) putative membrane protein [Streptomyces coelicolor A3(2)] | | synthase | 264605 |
| 918 | 91228795 (1835, 1836) | Novel Protein sim. GBank gi 168448 sp Q05813 JAMP1_STRLI - XAA-PRO AMINOPEPTIDASE I (X-PRO AMINOPEPTIDASE I) (AMINOPEPTIDASE P I) (APP) (PEPP I) | | | 264259, 28331826, 264908, 265019, 264448, 265020, 284635, 83373044 |
| 919 | 80436785 (1837, 1838) | Novel Protein sim. GBank gi 3850084 emb CAA21811.1 - (AL033389) alcohol dehydrogenase [Schizosaccharomyces pombe] | | | 265006, 264512, 264600, 264602, 264604, 264768, 18108370, 264563 |
| 920 | 79608095 (1839, 1840) | Novel Protein sim. GBank gi 1168448 sp Q05813 JAMP1_STRLI - XAA-PRO AMINOPEPTIDASE I (X-PRO AMINOPEPTIDASE I) (AMINOPEPTIDASE P I) (APP) (PEPP I) | | peptidase | 264508 |
| 921 | 19858634 (1841, 1842) | Novel Protein sim. GBank gi 3850084 emb CAA21811.1 - (AL033389) alcohol dehydrogenase [Schizosaccharomyces pombe] | | UNCLASSIFIED | 264600 |
| 922 | 78982605 (1843, 1844) | Novel Protein sim. GBank gi 267079 sp P29514 TB86_ARATH - TUBULIN BETA-6 CHAIN | | | 265018, 22279002 |
| 923 | 86695830 (1845, 1846) | Novel Protein sim. GBank gi 2497688 sp Q060963 PAFA_MOUSE - PLATELET- ACTIVATING FACTOR ACETYLHYDROLASE PRECURSOR (PAF ACETYLHYDROLASE) (PAF 2- ACYLHYDROLASE) (LDL-ASSOCIATED PHOSPHOLIPASE A2) (LDL-PLA(2)) (2-ACETYL-1- ALKYLGLYCEROPHOSPHOCHOLINE ESTERASE) (1- ALKYL-2-ACETYLGLYCEROPHO... | | tubulin | 264907, 265008, 265009, 265010, 18108351, 264689, 265021, 18108370, 18108374, 18108385 |
| 924 | 21431341 (1847, 1848) | Novel Protein sim. GBank gi 2497688 sp Q060963 PAFA_MOUSE - PLATELET- ACTIVATING FACTOR ACETYLHYDROLASE PRECURSOR (PAF ACETYLHYDROLASE) (PAF 2- ACYLHYDROLASE) (LDL-ASSOCIATED PHOSPHOLIPASE A2) (LDL-PLA(2)) (2-ACETYL-1- ALKYLGLYCEROPHOSPHOCHOLINE ESTERASE) (1- ALKYL-2-ACETYLGLYCEROPHO... | | esterase | 264510 |
| 925 | 20630332 (1849, 1850) | Novel Protein sim. GBank gi 3882325 dbj BAA34522.1 - (AB018345) KIAA0802 protein [Homo sapiens] | | | 264603 |
| 926 | 79397657 (1851, 1852) | Novel Protein sim. GBank gi 3882325 dbj BAA34522.1 - (AB018345) KIAA0802 protein [Homo sapiens] | | | 55811957, 263972, 264639 |
| 927 | 37036201 (1853, 1854) | | | UNCLASSIFIED | 264769 |

| | | | | | |
|-----|-----------------------|--|---|---------------|--|
| 928 | 80070610 (1855, 1856) | Novel Protein sim. GBank gij156146 (M30316) - xanthine dehydrogenase (AA at 2538) [Calliphora vicina] | | dehydrogenase | 264605 |
| 929 | 20630336 (1857, 1858) | | | UNCLASSIFIED | 264603 |
| 930 | 5496348 (1859, 1860) | Novel Protein sim. GBank gij115936[jgb]AAD03446.1] - (AF118223) No definition line found [Arabidopsis thaliana] | | | 264259 |
| 931 | 10245731 (1861, 1862) | Novel Protein sim. GBank gij4490809[emb]CAB38642.1] - (AJ133495) ribonucleotide reductase major subunit [Staphylococcus aureus] | Contains protein domain (PF00317) - Ribonucleotide reductase | reductase | 264486 |
| 932 | 80420613 (1863, 1864) | Novel Protein sim. GBank gij5459396[emb]CAB50754.1] - (AL096839) putative integral membrane transport protein [Streptomyces coelicolor] | | transport | 264508, 264905, 264906, 264909, 264800, 264602, 264603, 264605, 264766, 264558, 18108387, 264486 |
| 933 | 94326010 (1865, 1866) | Novel Protein sim. GBank gij5689523[dbj]BAA83045.1] - (AB029016) KIAA1093 protein [Homo sapiens] | Contains protein domain (PF00627) - UBA domain | UNCLASSIFIED | 264508, 264686, 264693, 27486261, 18108370, 65274791, 264636, 264559, 22279002 |
| 934 | 80039105 (1867, 1868) | Novel Protein sim. GBank gij119111[sp]P12978[EBN2_EBV - EBNA-2 NUCLEAR PROTEIN] | | | 264369, 22279002 |
| 935 | 80063162 (1869, 1870) | | | UNCLASSIFIED | 264604, 264605, 264693, 18108370, 18108374 |
| 936 | 80026632 (1871, 1872) | Novel Protein sim. GBank gij845686 (M32103) - ORF-27 [Staphylococcus aureus] | | UNCLASSIFIED | 22278995, 22278996, 264602, 264687, 32833986, 18108387 |
| 937 | 80250273 (1873, 1874) | Novel Protein sim. GBank gij1360669[pir]CGHU1V - collagen alpha 1(V) chain precursor - human | | synthase | 263978 |
| 938 | 80026633 (1875, 1876) | Novel Protein sim. GBank gij2226243[emb]CAA74531.1] - (Y14083) hypothetical protein [Bacillus subtilis] | | UNCLASSIFIED | 264602 |
| 939 | 11071694 (1877, 1878) | | | | 264600 |
| 940 | 94144252 (1879, 1880) | Novel Protein sim. GBank gij3560166[emb]CAA20678] - (AL031525) ubiquitin carboxyl-terminal hydrolase [Schizosaccharomyces pombe] | Contains protein domain (PF00442) - Ubiquitin carboxyl-terminal hydrolases family 2 | ubiquitin | 264905, 264906, 264907, 264908, 264909, 264511, 264910, 264592, 33657402, 264596, 264758, 264760, 264683, 264768, 264768, 264769, 33657023, 33657109, 264628, 264629, 264630, 264635, 264636, 264555, 264637, 264556, 264638, 264639, 83373044, 18108385, 264565, 18108391 |
| 941 | 11398414 (1881, 1882) | | | | 264593 |
| 942 | 19484122 (1883, 1884) | | | UNCLASSIFIED | 264760 |
| 943 | 80080258 (1885, 1886) | Novel Protein sim. GBank gij4033729 (AF038595) - apolipoprotein N-acyltransferase [Pseudomonas aeruginosa] | | UNCLASSIFIED | 264600, 264687, 264689, 264563 |
| 944 | 80216096 (1887, 1888) | Novel Protein sim. GBank gij2494764[sp]Q50729[GUAA_MYCTU - GMP SYNTHASE (GLUTAMINE-HYDROLYZING) (GLUTAMINE AMIDOTRANSFERASE) (GMP SYNTHETASE)] | | synthase | 264511, 264603 |
| 945 | 80052477 (1889, 1890) | Novel Protein sim. GBank gij732353[sp]P39606[YVCH_BACSU - HYPOTHETICAL 36.6 KD PROTEIN IN QOXD-VPR INTERGENIC REGION] | Contains protein domain (PF00818) - Ice nucleation protein repeat | UNCLASSIFIED | 264906, 264604, 264605, 265020, 18108387 |
| 946 | 78248402 (1891, 1892) | | | | 265017 |

| | | | | | |
|-----|-----------------------|--|---|----------------------|--|
| 947 | 81802698 (1893, 1894) | Novel Protein sim. GBank gij2896770[embjCAA17247] - (AL021899) hypothetical protein Rv2033c [Mycobacterium tuberculosis] | Contains protein domain (PF00459) - Inositol monophosphatase family | phosphatase | 18108394, 22276896, 264907, 264909, 265008, 265009, 264910, 264758, 264600, 264602, 265018, 264805, 264769, 264689, 264693 |
| 948 | 88185538 (1895, 1896) | Novel Protein sim. GBank gij2827284 (AF041037) - novel antagonist of FGF signaling [Homo sapiens] | | lgf | 18108398, 56182575, 22278997, 22278999, 60432049, 28331822, 29331826, 284907, 56182435, 55811386, 265011, 264600, 265017, 265018, 265019, 18108351, 265020, 265021, 265022, 27486265, 263972, 55811576, 264638, 60170394, 264566 |
| 949 | 88081786 (1897, 1898) | Novel Protein sim. GBank gij4507985[refjNP_003427.1 pZNF1 - zinc finger protein 135 (clone pHZ-17)] | Contains protein domain (PF00096) - Zinc finger, C2H2 type | transcription factor | 29331825, 21906764, 27486261, 21906766, 52644298, 33657349, 87188518, 56994075, 265020, 265021, 87188559, 52644150, 264637 |
| 950 | 79485872 (1899, 1900) | Novel Protein sim. GBank gij1079461[pij S43865 - cytochrome b, type II - potaroo (fragment)] | Contains protein domain (PF00038) - Intermediate filament proteins | struct | 264683, 18108361 |
| 951 | 20451411 (1901, 1902) | Novel Protein sim. GBank gij5420387[embjCAB46679.1] - (AJ243459) proteophosphoglycan [Leishmania major] | | UNCLASSIFIED | 264604 |
| 952 | 79566954 (1903, 1904) | Novel Protein sim. GBank gij5305702[gbjAAD41779.1 AF12686 - (AF126867) calpain-like protease [Mus musculus] | | cathepsin | 264910, 264691 |
| 953 | 10196003 (1905, 1906) | Novel Protein sim. GBank gij2495842[spjQ47142 YFHS_ECOLI - HYPOTHETICAL 41.8 KD PROTEIN IN CSIE-GLYA INTERGENIC REGION] | | transport | 264510 |
| 954 | 9893326 (1907, 1908) | Novel Protein sim. GBank gij2360985 (AF016253) - D-amino acid dehydrogenase [Klebsiella aerogenes] | | dehydrogenase | 264508 |
| 955 | 95313410 (1909, 1910) | Novel Protein sim. GBank gij5454064[refjNP_006319.1 pSIP] - SYT interacting protein | Contains protein domain (PF00076) - RNA recognition motif. (a.k.a. RRM, RBD, or RNP domain) | dna_rna_bind | 56994075, 264509, 264905, 264906, 264907, 264908, 264909, 264510, 264910, 264758, 264759, 265010, 264601, 264760, 18108351, 264762, 264763, 264764, 264766, 264686, 264767, 264687, 264768, 264769, 264689, 264628, 264629, 264630, 264631, 264632, 264634, 264635, 264636, 264637, 264638, 56182323, 264639, 18108388, 264563, 264564 |
| 956 | 80064224 (1911, 1912) | Novel Protein sim. GBank gij2052129[embjCAB08155] - (Z94752) rimJ [Mycobacterium tuberculosis] | | | 264805 |
| 957 | 80056208 (1913, 1914) | Novel Protein sim. GBank gij1709787[spjQ00451 PRF1_L_YCES - 36.4 KD PROLINE-RICH PROTEIN] | | UNCLASSIFIED | 264803, 18108362 |
| 958 | 80036446 (1915, 1916) | Novel Protein sim. GBank gij2131050[embjCAB09260] - (Z95844) opca [Mycobacterium tuberculosis] | | UNCLASSIFIED | 264908, 264910, 264762, 263978, 264637 |
| 959 | 80026647 (1917, 1918) | Novel Protein sim. GBank gij2128478[pij S51939 - chitinase (EC 3.2.1.14) precursor - beet] | | UNCLASSIFIED | 264602, 264692 |
| 960 | 37815406 (1919, 1920) | | | UNCLASSIFIED | 264259 |
| 961 | 20567383 (1921, 1922) | | | | 263978 |
| 962 | 11395318 (1923, 1924) | | | | 264593 |

| | | | | | | |
|-----|-----------------------|---|--|--|--------------|--|
| 963 | 80590374 (1925, 1926) | | | | UNCLASSIFIED | 264510, 264288, 264555, 264556, 264559, 264486 |
| 964 | 79832019 (1927, 1928) | Novel Protein sim. GBank gll4586622[db][BAA76833.1] - (AB023206) KIAA0889 protein [Homo sapiens] | | | UNCLASSIFIED | 264112, 264910, 264689 |
| 965 | 91229485 (1929, 1930) | Novel Protein sim. GBank gll542038[emb][CAB46679.1] - (AJ243459) proteophosphoglycan [Leishmania major] | | | UNCLASSIFIED | 264488, 265017, 264448, 264634, 264558, 83373044 |
| 966 | 95292815 (1931, 1932) | | | | UNCLASSIFIED | 264908, 264592, 264596, 264604, 264768, 21906764, 264692, 264693, 264629, 264636, 264638 |
| 967 | 79255708 (1933, 1934) | Novel Protein sim. GBank gll1731207[sp][Q11156]RGX3_MYCTU - SENSORY TRANSDUCTION PROTEIN REGX3 | Contains protein domain (PF00072) - Response regulator receiver domain | | phosphatase | 264780 |
| 968 | 78560269 (1935, 1936) | Novel Protein sim. GBank gll2661836[emb][CAA75187] - (Y14864) putative transport protein [Methylophilus methylotrophus] | | | transport | 264693 |
| 969 | 79918470 (1937, 1938) | Novel Protein sim. GBank gll5419878[emb][CAB46422.1] - (AL098747) hypothetical protein [Homo sapiens] | Contains protein domain (PF00098) - Zinc finger, C2H2 type | | dna_rna_bind | 35696286, 264685, 264686, 35695917, 264692, 18108374, 264635 |
| 970 | 95085947 (1939, 1940) | | | | UNCLASSIFIED | 18108392, 18108394, 18108398, 22278995, 22278996, 22278998, 22278999, 29147620, 264828, 265006, 265007, 265008, 265009, 18108348, 33109954, 265010, 265011, 18108351, 264288, 21908767, 21906768, 18108370, 18108374, 18108377, 264630, 264635, 18108380, 83373044, 18108387, 18108388 |
| 971 | 78919770 (1941, 1942) | | | | UNCLASSIFIED | 265007, 265020, 22279002 |
| 972 | 20710704 (1943, 1944) | | | | | 264557 |
| 973 | 20370183 (1945, 1946) | Novel Protein sim. GBank gll1723119[sp][P53990]Y174_HUMAN - HYPOTHETICAL PROTEIN KIAA0174 | | | | 264604 |
| 974 | 80037103 (1947, 1948) | | | | UNCLASSIFIED | 264565 |
| 975 | 10196018 (1949, 1950) | | | | UNCLASSIFIED | 264510 |
| 976 | 80205742 (1951, 1952) | Novel Protein sim. GBank gll3881459[emb][CAA92988.1] - (Z68753) predicted using GeneFinder; Similarity to Yeast hypothetical protein YIK9 (SW:YIK9_YEAST); cDNA EST EMBL:D27680 comes from this gene; cDNA EST EMBL:D27679 comes from this gene; cDNA EST EMBL:D64477 comes from this ge... | | | UNCLASSIFIED | 264508, 264906, 264758, 264632, 264639, 264563 |
| 977 | 10355349 (1953, 1954) | Novel Protein sim. GBank gll549456[sp][Q05335]XYS3_PSEPU - XYLDLEGF OPERON TRANSCRIPTIONAL ACTIVATOR 3 | | | UNCLASSIFIED | 264906 |
| 978 | 80025927 (1955, 1956) | | | | UNCLASSIFIED | 264600, 264602, 264603, 264604 |
| 979 | 80447820 (1957, 1958) | Novel Protein sim. GBank gll3171904[emb][CAA75869] - (Y15908) DIA-12C protein [Homo sapiens] | | | UNCLASSIFIED | 264767, 264768, 265006, 265007, 264906 |
| 980 | 80025928 (1959, 1960) | | | | UNCLASSIFIED | 264600, 264602, 264605 |
| 981 | 80098550 (1961, 1962) | Novel Protein sim. GBank gll3599940 (AF017368) - (adogenital dysplasia protein 2 [Mus musculus]) | | | UNCLASSIFIED | 264692, 264555, 264556, 264557, 264559 |

| | | | | |
|-----|-----------------------|--|--------------|--|
| 982 | 80195670 (1983, 1984) | Novel Protein sim. GBank gij2950220[emb]CAA71575] - (Y10545) fused-cdb [Escherichia coli] | UNCLASSIFIED | 264404 |
| 983 | 90995041 (1965, 1966) | Novel Protein sim. GBank gij476389[pil]B43402 - myosin heavy chain-B, neuronal - chicken | struct | 65274572, 56182575, 264908, 264909, 265007, 265008, 264758, 265010, 55811150, 33657023, 264634, 264557, 264558 |
| 984 | 20468878 (1987, 1988) | Novel Protein sim. GBank gij3451504[emb]CAA07660.1] - (A007747) hypothetical protein BbLPS1.21 [Bordetella bronchiseptica] | UNCLASSIFIED | 264605 |
| 985 | 85461368 (1969, 1970) | Novel Protein sim. GBank gij2105049[emb]CA088635] - (Z95436) hypothetical protein Rv3645 [Mycobacterium tuberculosis] | transferase | 56182435, 264600 |
| 986 | 87102868 (1971, 1972) | Novel Protein sim. GBank gij2105049[emb]CA088635] - (Z95436) hypothetical protein Rv3645 [Mycobacterium tuberculosis] | UNCLASSIFIED | 264106, 264110, 265020, 60170615 |
| 987 | 79867231 (1973, 1974) | Novel Protein sim. GBank gij2105049[emb]CA088635] - (Z95436) hypothetical protein Rv3645 [Mycobacterium tuberculosis] | UNCLASSIFIED | 264909 |
| 988 | 19858661 (1975, 1976) | Novel Protein sim. GBank gij2105049[emb]CA088635] - (Z95436) hypothetical protein Rv3645 [Mycobacterium tuberculosis] | UNCLASSIFIED | 264600 |
| 989 | 88095329 (1977, 1978) | Novel Protein sim. GBank gij2105049[emb]CA088635] - (Z95436) hypothetical protein Rv3645 [Mycobacterium tuberculosis] | UNCLASSIFIED | 264508, 265017, 264534, 264564 |
| 990 | 88057746 (1978, 1980) | Novel Protein sim. GBank gij2105049[emb]CA088635] - (Z95436) hypothetical protein Rv3645 [Mycobacterium tuberculosis] | UNCLASSIFIED | 264259, 264908, 265009, 264910, 264596, 264369, 264288, 264760, 264628, 264635, 264588 |
| 991 | 10108140 (1981, 1982) | Novel Protein sim. GBank gij2105049[emb]CA088635] - (Z95436) hypothetical protein Rv3645 [Mycobacterium tuberculosis] | UNCLASSIFIED | 264909 |
| 992 | 78845694 (1983, 1984) | Novel Protein sim. GBank gij2105049[emb]CA088635] - (Z95436) hypothetical protein Rv3645 [Mycobacterium tuberculosis] | UNCLASSIFIED | 264508, 264593 |
| 993 | 10814053 (1985, 1986) | Novel Protein sim. GBank gij2105049[emb]CA088635] - (Z95436) hypothetical protein Rv3645 [Mycobacterium tuberculosis] | UNCLASSIFIED | 264907 |
| 994 | 11090590 (1987, 1988) | Novel Protein sim. GBank gij2105049[emb]CA088635] - (Z95436) hypothetical protein Rv3645 [Mycobacterium tuberculosis] | reductase | 264602 |
| 995 | 94321911 (1988, 1990) | Novel Protein sim. GBank gij2105049[emb]CA088635] - (Z95436) hypothetical protein Rv3645 [Mycobacterium tuberculosis] | UNCLASSIFIED | 18108398, 65274572, 22278996, 264490, 60432049, 29331827, 29146498, 264508, 264905, 264907, 264908, 56182435, 265008, 264591, 264592, 60432229, 60431735, 33657402, 264595, 264758, 21906754, 265010, 265017, 265018, 264605, 264760, 264448, 264763, 264768, 21906785, 21906788, 21906789, 55811957, 264692, 264693, 264629, 35696423, 55811576, 35695855, 264636, 264555, 264556, 264558, 83373044, 22279002, 264563 |
| 996 | 91013745 (1991, 1992) | Novel Protein sim. GBank gij2911719 (AC004227) - KIA0011LB [Homo sapiens] | phosphatase | 83373044, 29331824, 263978, 55811957, 56528486, 87168518, 264910, 264908, 264565, 264566, 264693, 264766 |
| 997 | 80503347 (1993, 1994) | Novel Protein sim. GBank gij2649101 (AE001001) - ABC transporter, ATP-binding protein [Archaeoglobus fulgidus] | transport | 35696286, 22278997, 22278999, 264508, 264905, 264908, 265010, 264600, 264602, 264605, 264688, 264769, 265021, 264585, 18108391 |
| 998 | 11397390 (1995, 1996) | Novel Protein sim. GBank gij1235303[PD4928]HRPX_PLALO - HISTIDINE-RICH GLYCOPROTEIN PRECURSOR | | 264595 |

| | | | | | | |
|------|-----------------------|--|---|---------------|--|--|
| 999 | 11768047 (1997, 1998) | Novel Protein sim. GBank gi 2506697 sp P46480 YFCA_HAEIN - HYPOTHETICAL PROTEIN HI0198 | | UNCLASSIFIED | 264682 | |
| 1000 | 20727844 (1999, 2000) | | | UNCLASSIFIED | 264602 | |
| 1001 | 86673131 (2001, 2002) | Novel Protein sim. GBank gi 2224699 dbj BAA20833 - (AB002377) KIAA0379 [Homo sapiens] | Contains protein domain (PF00023) - Ank repeat | kinase | 60432049, 264807, 264909, 264511, 264603, 264693, 264684, 264687, 264689, 29148627, 21906769, 264692, 18108385, 22279000 | |
| 1002 | 80189603 (2003, 2004) | Novel Protein sim. GBank gi 586121 sp P37709 TRHY_RABIT - TRICHOHYALIN | | stud | 265009, 264369, 265020 | |
| 1003 | 17633491 (2005, 2006) | | | | 265019 | |
| 1004 | 18314887 (2007, 2008) | Novel Protein sim. GBank gi 854065 emb CAA58337 - (X93413) U88 [Human herpesvirus 6] | | | 264635 | |
| 1005 | 79617144 (2009, 2010) | Novel Protein sim. GBank gi 114073 sp P07672 APT_ECOLI - ADENINE PHOSPHORIBOSYLTRANSFERASE (APRT) | | UNCLASSIFIED | 264508 | |
| 1006 | 37815429 (2011, 2012) | | | UNCLASSIFIED | 264259 | |
| 1007 | 79620871 (2013, 2014) | Novel Protein sim. GBank gi 4062979 dbj BAA36210.1 - (AB017138) epsilon subunit of malonate decarboxylase [Pseudomonas putida] | | synthase | 264905 | |
| 1008 | 88094444 (2015, 2016) | Novel Protein sim. GBank gi 2808807 emb CAA04607.1 - (AJ001206) putative trehalose synthase [Streptomyces coelicolor] | | synthase | 265007, 264802, 264605, 264760, 264636 | |
| 1009 | 57451289 (2017, 2018) | Novel Protein sim. GBank gi 3639077 AF090113 - AMPA receptor binding protein [Rattus norvegicus] | Contains protein domain (PF00595) - POZ domain (Also known as DHR or GLGF). | kinase | 264102, 264288 | |
| 1010 | 94672537 (2019, 2020) | Novel Protein sim. GBank gi 3746332 AF016307 - possible NADH-dependent oxidase, may function as a demethylase [Sinorhizobium meliloti] | | dehydrogenase | 264592 | |
| 1011 | 85546918 (2021, 2022) | Novel Protein sim. GBank gi 2342647 gb AA86591.1 - (U90653) DHC-domain-containing cysteine-rich protein [Homo sapiens] | | UNCLASSIFIED | 35696052, 264905, 264764, 264768, 35695917, 264629 | |
| 1012 | 95294456 (2023, 2024) | Novel Protein sim. GBank gi 3413411 emb CAA20272 - (AL031231) guanosine pentaphosphate synthetase/ polyribonucleotide nucleotidyltransferase [Streptomyces coelicolor] | Contains protein domain (PF00013) - KH domain | phosphorylase | 35696052, 264905, 264600, 264601, 264602, 264605, 264762, 264766, 264768, 264689 | |
| 1013 | 86095772 (2025, 2026) | | | UNCLASSIFIED | 264591, 21906768 | |
| 1014 | 86608828 (2027, 2028) | | | | 29331824, 265019, 265020 | |

| | | | | | |
|------|-----------------------|---|---|-------------------|--|
| 1015 | 95418879 (2029, 2030) | Novel Protein sim. GBank gi 4159995 (AF063095) - SELIL [Mus musculus] | Contains protein domain (PF00040) - Fibronectin type II domain | struct | 22278994, 22278995, 56994075, 22278996, 22278999, 264259, 29331825, 29331828, 264907, 56182435, 264510, 264591, 264593, 60433356, 264594, 55812038, 264758, 21908754, 33857084, 265010, 264600, 265017, 265018, 265019, 18108351, 21906765, 21906766, 21906767, 21906768, 55811957, 265022, 33657023, 65274620, 33657182, 32833986, 18108370, 18108377, 55811576, 35696423, 264630, 22279000, 264565 |
| 1016 | 79559694 (2031, 2032) | Novel Protein sim. GBank gi 25069899 p41407 ACPD_ECOLI - ACYL CARRIER PROTEIN PHOSPHODIESTERASE (ACP PHOSPHODIESTERASE) | | esterase | 264686, 264693 |
| 1017 | 11069213 (2033, 2034) | Novel Protein sim. GBank gi 5103943 dbj BAA79259.1 - (AP000059) 802aa long hypothetical oligopeptide-binding protein oppA [Aeropyrum pernix] | Contains protein domain (PF00496) - Bacterial extracellular solute-binding proteins, family 5 | transport | 264600 |
| 1018 | 80072430 (2035, 2036) | Novel Protein sim. GBank gi 4493973 emb CAB39032.1 - (AL034559) predicted using hexExon; MAL3P7.14 (PFC0925w). Hypothetical protein, len: 489 aa [Plasmodium falciparum] | | | 22278996, 29148627, 264563 |
| 1019 | 11703607 (2037, 2038) | | | UNCLASSIFIED | 264686 |
| 1020 | 80234432 (2039, 2040) | | | | 264508, 264509, 264512, 264600, 264762, 264769, 264688, 18108370, 264636, 264638, 264486 |
| 1021 | 37036243 (2041, 2042) | Novel Protein sim. GBank gi 4633807 gb AAD26859.1 AF12779 - (AF127795) trehalose biosynthetic enzyme TreY [Rhizobium leguminosarum bv. viciae] | | synthase | 264769 |
| 1022 | 80502627 (2043, 2044) | Novel Protein sim. GBank gi 1781230 emb CAB06271 - (Z83867) hypothetical protein Rv3137 [Mycobacterium tuberculosis] | Contains protein domain (PF00459) - Inositol monophosphatase family | phosphatase | 35696052, 264508, 265008, 265009, 264769, 18108387, 264563 |
| 1023 | 11399341 (2045, 2046) | Novel Protein sim. GBank gi 3777495 (U92083) - calcium transporting ATPase [Pichia angusta] | Contains protein domain (PF00122) - E1-E2 ATPase | ATPase-associated | 264593 |
| 1024 | 80057129 (2047, 2048) | | | UNCLASSIFIED | 52646842, 33657402, 33657023, 18108379, 55811576, 264631, 264556, 264557, 264559, 18108388, 264566 |
| 1025 | 79644200 (2049, 2050) | Novel Protein sim. GBank gi 3483045 emb CAA20556 - (AL031371) putative transport system permease protein [Streptomyces coelicolor] | | transport | 264693 |
| 1026 | 80025948 (2051, 2052) | Novel Protein sim. GBank gi 1174922 sp Q02322 LUVRD - HAEIN - DNA HELICASE II | | helicase | 264602 |
| 1027 | 17659234 (2053, 2054) | Novel Protein sim. GBank gi 4757728 ref NP_004886.1 pAGTA - angiotensin/vasopressin receptor AII/AVP-like | | UNCLASSIFIED | 265017 |

| | | | | | |
|------|-----------------------|---|--|--------------|--|
| 1028 | 20297928 (2055, 2056) | Novel Protein sim. GBank gij2791409[emb][CAA16003] - (AL021184) aco [Mycobacterium tuberculosis] | Contains protein domain (PF00330) - Aconitase family (aconitate hydratase) | UNCLASSIFIED | 264800 |
| 1029 | 94665090 (2037, 2058) | | | UNCLASSIFIED | 264595 |
| 1030 | 88095343 (2059, 2060) | | | UNCLASSIFIED | 264907, 264908, 264510, 264512, 265009, 265010, 265011, 264600, 264602, 264603, 264605, 264769, 18108372, 18108374 |
| 1031 | 95289117 (2061, 2062) | | | UNCLASSIFIED | 264905, 264906, 264909, 264595, 264692, 264630, 264634, 264636 |
| 1032 | 94673275 (2063, 2064) | Novel Protein sim. GBank gij4503895[ref]NP_000145.1pGALK - galactokinase 1 | | UNCLASSIFIED | 264689 |
| 1033 | 86464818 (2065, 2066) | Novel Protein sim. GBank gij2982990 (AE000682) - hypothetical protein [Aquifex aeolicus] | | UNCLASSIFIED | 35896052, 264908, 264510, 18108354, 264687, 264769, 264689, 80431602, 18108385, 264486 |
| 1034 | 79245937 (2067, 2068) | Novel Protein sim. GBank gij405895 (U00007) - methionyl-tRNA synthetase [Escherichia coli] | | UNCLASSIFIED | 264906 |
| 1035 | 79956355 (2069, 2070) | | | UNCLASSIFIED | 264692 |
| 1036 | 85804998 (2071, 2072) | | | UNCLASSIFIED | 264905, 66712502, 264908, 264766 |
| 1037 | 8786058 (2073, 2074) | | | UNCLASSIFIED | 29331824, 264809, 80433438, 265019 |
| 1038 | 20481015 (2075, 2076) | Novel Protein sim. GBank gij790819 (L39891) - polycystic kidney disease-associated protein [Homo sapiens] | Contains protein domain (PF01477) - PLAT/LH2 domain | | 264604, 264634 |
| 1039 | 87260021 (2077, 2078) | Novel Protein sim. GBank gij2605967 (AF030027) - 24 [Equine herpesvirus 4] | | UNCLASSIFIED | 264092, 264093, 264094, 264683, 264689, 263967 |
| 1040 | 80026840 (2079, 2080) | Novel Protein sim. GBank gij2352095 (U97022) - DNA topoisomerase I [Fervidobacterium islandicum] | Contains protein domain (PF01131) - Prokaryotic DNA topoisomerase | isomerase | 264595 |
| 1041 | 10156682 (2081, 2082) | Novel Protein sim. GBank gij3256535[dbj]BAA29218.1 - (AP000001) 301aa long hypothetical 2-phosphoglycerate kinase [Pyrococcus horikoshii] | | kinase | 264907 |
| 1042 | 11084375 (2083, 2084) | Novel Protein sim. GBank gij2058299[emb][CAA66953] - (X98309) ARI protein [Drosophila melanogaster] | | | 264605 |
| 1043 | 80057136 (2085, 2086) | Novel Protein sim. GBank gij1870167[emb][CAA70125] - (Y08921) msK [Streptomyces reticuli] | Contains protein domain (PF00005) - ABC transporter | transport | 264565, 264567 |
| 1044 | 80025952 (2087, 2088) | Novel Protein sim. GBank gij5689890[emb][CAB52053.1] - (AL109732) hypothetical protein [Streptomyces coelicolor A3(2)] | | UNCLASSIFIED | 265006, 264602, 265017 |
| 1045 | 52415482 (2089, 2090) | | | | 29331825, 264637 |
| 1046 | 11754862 (2091, 2092) | Novel Protein sim. GBank gij854065[emb][CAA58337] - (X83413) U88 [Human herpesvirus 6] | | helicase | 264686 |
| 1047 | 37036238 (2093, 2094) | Novel Protein sim. GBank gij4210471[dbj]BAA74535.1 - (AB019033) orfSA [Pseudomonas sp.] | | UNCLASSIFIED | 264769 |
| 1048 | 79186400 (2095, 2096) | Novel Protein sim. GBank gij3413419[emb][CAA20279] - (AL031232) hypothetical protein SC10H5.07 [Streptomyces coelicolor] | Contains protein domain (PF00220) - Neurohypophysial hormones, N-terminal Domain | UNCLASSIFIED | 264687 |
| 1049 | 81755108 (2097, 2098) | Novel Protein sim. GBank gij5051636[gb]AAD38328.1[AF073727] EH domain-binding mitotic phosphoprotein [Homo sapiens] | | UNCLASSIFIED | 264905, 264634 |
| 1050 | 79471521 (2099, 2100) | | | UNCLASSIFIED | 264686 |

| | | | | | | |
|------|-----------------------|---|--|--|------------------|--|
| 1051 | 80475471 (2101, 2102) | | | | UNCLASSIFIED | 18108374, 264769, 265010, 265011, 264601, 265009, 264604, 264605, 264636, 18108351, 264692 |
| 1052 | 82442982 (2103, 2104) | Novel Protein sim. GBank gi 3123275 sp P35138 SERA_BACSU - D-3- PHOSPHOGLYCERATE DEHYDROGENASE (PGDH)- Novel Protein sim. GBank gi 5441319 emb CAB46717.1 - (AL034396) dJ1158812.1 (zinc finger, X-linked, duplicated A) [Homo sapiens] | Contains protein domain (PF00389) - D-isomer specific 2-hydroxyacid dehydrogenases | | dehydrogenase | 264508, 264762, 264687, 264486 |
| 1053 | 94851640 (2105, 2106) | | | | | 264686, 18108374, 29331824, 83373044, 21906754, 52645158, 56182435, 264689, 29331827, 27486281, 35696052, 21908765, 35696423, 21906768, 56182575, 21908769, 55811957, 87168518, 35696286, 22278997, 265020, 265011, 265021, 265022, 265007, 265018, 22279000, 22279002, 264482, 264906, 52644150, 264909, 264288, 29331822, 52645080, 264768 |
| 1054 | 79580225 (2107, 2108) | | | | UNCLASSIFIED | 264686 |
| 1055 | 80594138 (2109, 2110) | Novel Protein sim. GBank gi 5052508 gb AAD38584.1 AF145609) BCDNA_GH02833 [Drosophila melanogaster] Novel Protein sim. GBank gi 3021676 dbj BAA25358 - (D86033) RNA polymerase sigma-70 factor [Pseudomonas fluorescens] Novel Protein sim. GBank gi 1226281 (U50308) - No definition line found [Caenorhabditis elegans] | Contains protein domain (PF00270) - DEAD/DEAH box helicase | | helicase | 264907, 264602, 264681, 264288, 21906768, 33657109, 55810764, 35695855, 264631 |
| 1056 | 17882319 (2111, 2112) | | | | mapolymerase | 264906 |
| 1057 | 85667216 (2113, 2114) | | | | UNCLASSIFIED | 264682 |
| 1058 | 80376576 (2115, 2116) | | | | | 264764 |
| 1059 | 84662754 (2117, 2118) | Novel Protein sim. GBank gi 1170016 sp P46808 GREY_MYCLE - TRANSCRIPTION ELONGATION FACTOR GREY (TRANSCRIPT CLEAVAGE FACTOR GREY) Novel Protein sim. GBank gi 2499087 sp Q09332 UGGG_DROME - UDP- GLUCOSE:GLYCOPROTEIN GLUCOSYLTRANSFERASE PRECURSOR (DUGT) Novel Protein sim. GBank gi 90254 pir J28334 - protein- tyrosine-phosphatase (EC 3.1.3.48) Ly-5 precursor (B-cell variant) - mouse Novel Protein sim. GBank gi 3334200 sp O49954 GCSP_SOLTU - GLYCINE DEHYDROGENASE (DECARBOXYLATING) PRECURSOR (GLYCINE DECARBOXYLASE) (GLYCINE CLEAVAGE SYSTEM P-PROTEIN) Novel Protein sim. GBank gi 2499968 sp Q41228 PSE1_NICSY - PHOTOSYSTEM I REACTION CENTRE SUBUNIT IV A PRECURSOR (PSI-E A) Novel Protein sim. GBank gi 82654 pir JUA0086 - 10K zein precursor - maize | | | transcriptfactor | 35696052, 35695855, 265009, 264636 |
| 1060 | 79481169 (2119, 2120) | | | | glycoprotein | 29146498, 264681, 264683, 264687 |
| 1061 | 11034025 (2121, 2122) | | | | phosphatase | 264634 |
| 1062 | 39567837 (2123, 2124) | | | | dehydrogenase | 264593 |
| 1063 | 8480481 (2125, 2126) | | | | | 264508 |
| 1064 | 78891783 (2127, 2128) | | | | | 265007, 265008, 18108351, 18108385 |

| | | | | | |
|------|-----------------------|--|---|-------------------|--|
| 1065 | 80021208 (2129, 2130) | Novel Protein sim. GBank gij212098[pirt][S70682 - glycosyltransferase homolog - Bordetella pertussis] | | transferase | 264600, 264602, 264689 |
| 1066 | 17896879 (2131, 2132) | Novel Protein sim. GBank gij2508382[sp]P15042[DNLJ_ECOLI - DNA LIGASE (POLYDEOXYRIBONUCLEOTIDE SYNTHASE (NAD+))] | | synthase | 265009 |
| 1067 | 10132178 (2133, 2134) | | | | 264909 |
| 1068 | 82062057 (2135, 2136) | Novel Protein sim. GBank gij4007668[emb]CAA22355] - (AL034443) putative oxidoreductase [Streptomyces coelicolor] | Contains protein domain (PF00248) - Aldol/keto reductase family | reductase | 264688, 18108362, 264558, 264600, 264760 |
| 1069 | 83002954 (2137, 2138) | Novel Protein sim. GBank gij4589484[dbj]BAA76770.1] - (AB023143) KIAA0928 protein [Homo sapiens] | | UNCLASSIFIED | 264604 |
| 1070 | 82101992 (2139, 2140) | Novel Protein sim. GBank gij120304[sp]P15932[FLGK_SALTY - FLAGELLAR HOOK-ASSOCIATED PROTEIN 1 (HAP1)] | | UNCLASSIFIED | 264604, 264760 |
| 1071 | 20710589 (2141, 2142) | Novel Protein sim. GBank gij1750397 (U81261) - glutamate synthase large subunit [Pseudomonas aeruginosa] | | synthase | 264602 |
| 1072 | 82356540 (2143, 2144) | | | UNCLASSIFIED | 264687, 264688, 21908784, 35696052, 35695917, 35695855, 264600, 264601, 264602, 265009, 264605, 264508, 264905, 264690, 264906, 264762, 264628, 264768 |
| 1073 | 79814400 (2145, 2146) | Novel Protein sim. GBank gij477532[pirt]A49175 - Motch B protein - mouse (fragment) | Contains protein domain (PF00008) - EGF-like domain | synthase | 264909 |
| 1074 | 80105992 (2147, 2148) | Novel Protein sim. GBank gij3893109[emb]CAA76940] - (Y17920) CALO protein [Drosophila melanogaster] | | UNCLASSIFIED | 264908 |
| 1075 | 81850293 (2148, 2150) | | | ATPase-associated | 56994075, 22278998, 264594, 264757, 264596, 265018, 265019, 264681, 264369, 264688, 265020, 18108364, 18108374 |
| 1076 | 80477264 (2151, 2152) | Novel Protein sim. GBank gij1176203[sp]P46442[YHCM_ECOLI - HYPOTHETICAL 43.1 KD PROTEIN IN RPLM-HHOA INTERGENIC REGION (F375)] | | | 264769 |
| 1077 | 79831334 (2153, 2154) | Novel Protein sim. GBank gij4033487[sp]Q44472[TUD4_AGRVI - PUTATIVE HYDROXYPYRUVATE REDUCTASE] | | kinase | 264905 |
| 1078 | 20288874 (2155, 2156) | | | UNCLASSIFIED | 264600 |
| 1079 | 80494518 (2157, 2158) | Novel Protein sim. GBank gij3413828[emb]CAA20296] - (AL031260) hypothetical protein SC9A10.09 [Streptomyces coelicolor] | | UNCLASSIFIED | 18108394, 264769, 264634, 264636 |
| 1080 | 11767188 (2159, 2160) | | | UNCLASSIFIED | 264684 |
| 1081 | 94747080 (2161, 2162) | | | | 83373044, 265019, 22278002, 264482, 18108351, 264682, 264908, 264693, 264487 |
| 1082 | 81490656 (2163, 2164) | | | UNCLASSIFIED | 264758, 264768, 264769, 21906767, 264511, 264910, 264634, 264635, 264905, 264636, 264906, 264637, 264907, 264908, 264764, 264638, 20281089, 264768, 264595 |

| | | | | | |
|------|-----------------------|---|--|---------------|---|
| 1083 | 87446717 (2165, 2166) | Novel Protein sim. GBank gij1722845[sp]Q10523[Y01N_MYCTU - HYPOTHETICAL 44.6 KO PROTEIN CY427.23 | | UNCLASSIFIED | 60424179, 264905, 264908, 264510, 60432229, 264759, 87168474, 264605, 264769, 264689, 18108364, 18108376, 35695855, 264636 264769 |
| 1084 | 37789308 (2167, 2168) | Novel Protein sim. GBank gij18384[sp]P32057[WCAI_ECOLI - PUTATIVE COLANIC ACID BIOSYNTHESIS GLYCOSYL TRANSFERASE WCAI | | UNCLASSIFIED | |
| 1085 | 86475368 (2169, 2170) | Novel Protein sim. GBank gij1899190 (U90204) - heat shock protein 60 [Tsukamurella tyrosinosolvens] | Contains protein domain (PF00118) - TCP-1/cpn60 chaperonin family | eph | 60432229, 264687 |
| 1086 | 79608269 (2171, 2172) | Novel Protein sim. GBank gij1172858[sp]P46176[RL14_ACYKS - 50S RIBOSOMAL PROTEIN L14 | Contains protein domain (PF00238) - Ribosomal protein L14 | ribosomalprot | 264486 |
| 1087 | 78603978 (2173, 2174) | Novel Protein sim. GBank gij160199[emb]CAA15431 - (AL008583) dJ327J16.3 (novel CHROMobox family protein) [Homo sapiens] | Contains protein domain (PF00385) - chromo' (CHR)romatin Organization Modifier domain | helicase | 29331827, 264693 |
| 1088 | 78654963 (2175, 2176) | Novel Protein sim. GBank gij2983155 (AE000593) - phosphoglucosyltransferase/phosphomannomutase [Aquifex aeolicus] | | UNCLASSIFIED | 264905, 264601, 18108387 |
| 1089 | 80216800 (2177, 2178) | Novel Protein sim. GBank gij4981768[gb]AAD36290.1(AE00177) - (AE001778) NADH dehydrogenase, 30 kDa subunit, putative [Thermotoga maritima] | Contains protein domain (PF00329) - Respiratory-chain NADH dehydrogenase, 30 kDa subunit | UNCLASSIFIED | 264488, 264511, 265011, 264682, 264768, 264689, 21806764, 35695917, 265020, 32633986, 18108370, 35695855 |
| 1090 | 11083825 (2179, 2180) | Novel Protein sim. GBank gij4007680[emb]CAA22366 - (AL034443) putative oxidoreductase [Streptomyces coelicolor] | | | 264604 |
| 1091 | 12817471 (2181, 2182) | Novel Protein sim. GBank gij2495562[sp]P77239[YLCD_ECOLI - HYPOTHETICAL 44.3 KO PROTEIN IN NFRB-PHEP INTERGENIC REGION PRECURSOR | | UNCLASSIFIED | 264637 |
| 1092 | 80252286 (2183, 2184) | Novel Protein sim. GBank gij2960098[emb]CAA17996.1 - (AL022121) ntl [Mycobacterium tuberculosis] | Contains protein domain (PF00730) - Endonuclease III | nuclease | 264566 264769, 35695917, 35695855, 264600, 264602, 264603, 264605, 18108351 |
| 1093 | 80496304 (2185, 2186) | Novel Protein sim. GBank gij1001642[db]BAA10373 - (D84002) dGTP triphosphohydrolase [Synecococcus sp.] | | UNCLASSIFIED | 264686 |
| 1094 | 10880972 (2187, 2188) | Novel Protein sim. GBank gij4585587[emb]CAB40855.1 - (AL049828) putative adenine glycosylase [Streptomyces coelicolor] | Contains protein domain (PF00455) - Bacterial regulatory proteins, deoR family | nuclease | 264906, 265007, 264595, 264600, 264602, 264603, 264604, 264605, 264762, 264768, 264769, 264636, 264558, 18108387, 80432113, 284482, 264486 |
| 1095 | 87457250 (2189, 2190) | Novel Protein sim. GBank gij4585587[emb]CAB40855.1 - (AL049828) putative adenine glycosylase [Streptomyces coelicolor] | | | 264600, 264602, 264603, 264604, 264605, 35695917, 264692, 264631 |
| 1096 | 80025977 (2191, 2192) | Novel Protein sim. GBank gij115001[sp]P19206[BIOB_BACSH - BIOTIN SYNTHASE (BIOTIN SYNTHETASE)] | | synthase | |
| 1097 | 79239560 (2193, 2194) | Novel Protein sim. GBank gij114135[sp]P08205[JARGA_ECOLI - AMINO-ACID ACETYLTRANSFERASE (N-ACETYLGLUTAMATE SYNTHASE) (AGS)] | | UNCLASSIFIED | 285019 |
| 1098 | 79186424 (2195, 2196) | Novel Protein sim. GBank gij114135[sp]P08205[JARGA_ECOLI - AMINO-ACID ACETYLTRANSFERASE (N-ACETYLGLUTAMATE SYNTHASE) (AGS)] | | synthase | 264687 |

| | | | | | | |
|------|-----------------------|--|--|--|---------------|--|
| 1099 | 39523838 (2197, 2198) | Novel Protein sim. GBank gi 3915144 sp O33017 TRMD_MYCLE - TRNA (GUANINE- N1)-METHYLTRANSFERASE (MIG- METHYLTRANSFERASE) (TRNA [GM37] METHYLTRANSFERASE) | | | UNCLASSIFIED | 264603 |
| 1100 | 85736571 (2199, 2200) | Novel Protein sim. GBank gi 3023255 sp Q64420 ACOD_MESAU - ACYL-COA DESATURASE (STEAROYL-COA DESATURASE) (FATTY ACID DESATURASE) (DELTA(9)-DESATURASE) | | | desaturase | 264259, 264636 |
| 1101 | 80491857 (2201, 2202) | Novel Protein sim. GBank gi 1174735 sp P43012 TOP1_HAEIN - DNA TOPOISOMERASE I (OMEGA-PROTEIN) (RELAXING ENZYME) (UNTWISTING ENZYME) (SWIVELASE) | Contains protein domain (PF01396) - Topoisomerase DNA binding C4 zinc finger | | isomerase | 264769 |
| 1102 | 79777614 (2203, 2204) | Novel Protein sim. GBank gi 1906596 (U81788) - kinesin-73 | | | UNCLASSIFIED | 264910, 264909 |
| 1103 | 81897259 (2205, 2206) | [Drosophila melanogaster] | | | struct | 264757 |
| 1104 | 95003115 (2207, 2208) | Novel Protein sim. GBank gi 2935448 (AF048976) - synaptic ras GTPase-activating protein p135 SynGAP (Rattus norvegicus) | | | UNCLASSIFIED | 29331822, 21906754, 264555, 264556, 264558, 22279002 |
| 1105 | 80255121 (2209, 2210) | | | | | 264566 |
| 1106 | 79314110 (2211, 2212) | | | | UNCLASSIFIED | 264555, 264369 |
| 1107 | 80470019 (2213, 2214) | | | | | 264906, 264769 |
| 1108 | 80440616 (2215, 2216) | Novel Protein sim. GBank gi 1173421 sp P43416 SECY_STRSC - PREPROTEIN TRANSLOCASE SECY SUBUNIT | | | transport | 264907, 264510, 264511, 264600, 264602, 264605, 264768, 264769 |
| 1109 | 80064615 (2217, 2218) | Novel Protein sim. GBank gi 2895310 emb CAA18338 - (AL022268) putative ATP-dependent helicase (Streptomyces coelicolor) | | | helicase | 264602, 264605, 264636 |
| 1110 | 80503554 (2219, 2220) | | | | | 264908, 264593, 265010, 264601, 264603, 264604, 264605, 264682, 264769, 264693, 264636 |
| 1111 | 80071744 (2221, 2222) | Novel Protein sim. GBank gi 2622039 (AE000868) - type I restriction modification system, subunit S [Methanobacterium thermoautotrophicum] | | | | 18108370, 264557 |
| 1112 | 95010088 (2223, 2224) | | | | | 264908 |
| 1113 | 82456352 (2225, 2226) | Novel Protein sim. GBank gi 3218376 emb CAA18628 - (AL023862) putative oxidoreductase [Streptomyces coelicolor] | | | UNCLASSIFIED | 264600, 264602, 264604, 264605, 264762, 264769, 264565 |
| 1114 | 14998014 (2227, 2228) | Novel Protein sim. GBank gi 1083428 pir S54876 - NAD(P)+ transhydrogenase (B-specific) (EC 1.6.1.1) precursor - mouse | | | dehydrogenase | 264636 |
| 1115 | 11765583 (2229, 2230) | | | | | 264686 |
| 1116 | 79841152 (2231, 2232) | | | | UNCLASSIFIED | 264908 |

| | | | | | |
|------|-----------------------|--|---|--------------|--|
| 1117 | 95305485 (2233, 2234) | Novel Protein sim. GBank gi 3255965 emb CAA94089 - (Z70200) U5 snRNP-specific 200kD protein [Homo sapiens] | Contains protein domain (PF00270) - DEAD/DEAH box helicase | helicase | 18108392, 264488, 263994, 264489, 18108398, 56182575, 22278995, 22278996, 35696286, 56994075, 22278997, 22278998, 22278999, 264259, 29331822, 56182181, 29331824, 68714117, 29331825, 29331826, 60432289, 29331827, 29331828, 35696052, 33656970, 28146498, 284508, 264905, 264509, 264906, 264907, 264908, 66712502, 264909, 52644045, 56182435, 264510, 264511, 265006, 265007, 264512, 265008, 265009, 264910, 60170831, 264591, 264592, 264593, 60433356, 264594, 60433438, 264595, 55812038, 264596, 264758, 264759, 21906754, 33109954, 33657084, 265011, 87168559, 264600, 264601, 264602, 264603, 265017, 264604, 265018, 264605, 265019, 264760, 55811150, 264681, 264762, 18108351, 264448, 264682, 264764, 264683, 264288, 264389, 264684, 264685, 264766, 264767, 264686, 264687, 264768, 264769, 264688, 18108359, 264689, 21906765, 21906766, 21906767, 21906768, 21906769, 55811957, 29148784, 35695917, 265020, 265021, 265022, 60170615, 264690, 264691, 33657023, 264692, 264693, 65274620, 33657109, 33657182, 27486261, 27486262, 27486264, 33657349, 27486285, 35695763, 264691 |
| 1118 | 79563326 (2235, 2236) | | | UNCLASSIFIED | 264691 |
| 1119 | 79642463 (2237, 2238) | | | UNCLASSIFIED | 264907 |
| 1120 | 79480463 (2239, 2240) | Novel Protein sim. GBank gi 5420387 emb CAB46679.1 - (AJ243459) proteophosphoglycan [Leishmania major] | | collagen | 29331827, 265018, 265019, 264681, 265021, 60170615, 18108387 |
| 1121 | 79471716 (2241, 2242) | Novel Protein sim. GBank gi 1644450 U67864 - MEX-3 [Caenorhabditis elegans] | Contains protein domain (PF00013) - KH domain | UNCLASSIFIED | 264683, 264632, 18108388 |
| 1122 | 79456246 (2243, 2244) | | | UNCLASSIFIED | 264639, 264563 |
| 1123 | 79637119 (2245, 2246) | Novel Protein sim. GBank gi 98800 pir J17768 - 3-dehydroquinolate synthase (EC 4.6.1.3) - Mycobacterium tuberculosis | | synthase | 264683, 27486265 |
| 1124 | 79811596 (2247, 2248) | | | UNCLASSIFIED | 264909 |
| 1125 | 79757861 (2249, 2250) | | | UNCLASSIFIED | 264910 |
| 1126 | 79758914 (2251, 2252) | Novel Protein sim. GBank gi 138154 sp P03643 VGG_BPPHX - MAJOR SPIKE PROTEIN (G PROTEIN) (GPG) | | eph | 264905, 264909, 264910 |
| 1127 | 11800930 (2253, 2254) | | | UNCLASSIFIED | 264682 |
| 1128 | 8384885 (2255, 2256) | Novel Protein sim. GBank gi 5002704 emb CAB44358.1 - (AJ242630) DNA polymerase I [Methylobacterium sp. DMA] | Contains protein domain (PF00476) - DNA polymerase family A | polymerase | 264511 |

| | | | | | |
|------|-----------------------|--|--|--------------|--|
| 1129 | 80422480 (2257, 2258) | Novel Protein sim. GBank gi 5689485 dbj BAA03026.1 - (AB028997) KIAA1074 protein [Homo sapiens] | Contains protein domain (PF00170) - bZIP transcription factor | UNCLASSIFIED | 265011, 264766 |
| 1130 | 78420151 (2259, 2260) | Novel Protein sim. GBank gi 4981328 gb AAD35881.1 AE00174 - (AE001747) bioY protein [Thermotoga maritima] | | UNCLASSIFIED | 264595 |
| 1131 | 80055391 (2261, 2262) | Novel Protein sim. GBank gi 1841552 (U89336) - unknown [Homo sapiens] | | UNCLASSIFIED | 35696286, 22278998, 29331828, 264603, 264605, 264559 |
| 1132 | 82062248 (2263, 2264) | Novel Protein sim. GBank gi 1841552 (U89336) - unknown [Homo sapiens] | | UNCLASSIFIED | 22278996, 264906, 265009, 264600, 264602, 264604, 264605, 264760, 32833986, 18108374 |
| 1133 | 17280437 (2265, 2266) | | | UNCLASSIFIED | 265018 |
| 1134 | 80235376 (2267, 2268) | | | UNCLASSIFIED | 264512, 264534 |
| 1135 | 80029393 (2269, 2270) | Novel Protein sim. GBank gi 4539171 emb CAB39700.1 - (AL049485) conserved hypothetical protein [Streptomyces coelicolor] | | UNCLASSIFIED | 264508, 264600, 264602, 264603, 18108376 |
| 1136 | 79842052 (2271, 2272) | Novel Protein sim. GBank gi 4982454 gb AAD36931.1 AE00182 - (AE001823) ATP-dependent protease LA ₁ putative [Thermotoga maritima] | | UNCLASSIFIED | 264906, 264908 |
| 1137 | 90931557 (2273, 2274) | Novel Protein sim. GBank gi 4972746 gb AAD34768.1 - (AF132180) unknown [Drosophila melanogaster] | Contains protein domain (PF00515) - TPR Domain | collagen | 22278998, 22278999, 35696052, 264907, 265009, 60433396, 264598, 265010, 264448, 264882, 264767, 264689, 265020, 264692, 55811576, 35695855, 264631, 264632, 22278002 |
| 1138 | 79841163 (2275, 2276) | Novel Protein sim. GBank gi 731607 sp P38739 YHC8_YEAST - HYPOTHETICAL 63.8 KD PROTEIN IN GUT1-RIM1 INTERGENIC REGION PRECURSOR | | strud | 264908 |
| 1139 | 79633561 (2277, 2278) | Novel Protein sim. GBank gi 3650031 (AC005396) - putative proline-rich cell wall protein [Arabidopsis thaliana] | | UNCLASSIFIED | 264893 |
| 1140 | 39480358 (2279, 2280) | | | UNCLASSIFIED | 264593 |
| 1141 | 79638019 (2281, 2282) | | | UNCLASSIFIED | 265019, 264693 |
| 1142 | 19635848 (2283, 2284) | | | UNCLASSIFIED | 264631 |
| 1143 | 87762158 (2285, 2286) | Novel Protein sim. GBank gi 3928000 emb CAA05880 - (AJ003125) procollagen I N-proteinase [Homo sapiens] | Contains protein domain (PF00090) - Thrombospondin type 1 domain | oxidase | 56182575, 264908, 264600, 264632, 87168518 |
| 1144 | 80068998 (2287, 2288) | | | UNCLASSIFIED | 264635, 264636, 264907, 264593, 264908, 264566, 264909 |
| 1145 | 14610262 (2289, 2290) | | | UNCLASSIFIED | 264112 |
| 1146 | 82062092 (2291, 2292) | | | UNCLASSIFIED | 264769, 264689, 35696286, 264760, 264905, 264486, 264559 |
| 1147 | 80071761 (2293, 2294) | | | | 264557 |
| 1148 | 80048433 (2295, 2296) | Novel Protein sim. GBank gi 2499003 sp P76422 THID_ECOLI - PHOSPHOMETHYL PYRIMIDINE KINASE (HMP-PHOSPHATE KINASE) (HMP-P KINASE) | kinase | | 264591 |
| 1149 | 11607438 (2297, 2298) | Novel Protein sim. GBank gi 2896734 emb CAA17213.1 - (AL021897) hypothetical protein Rv1097c [Mycobacterium tuberculosis] | | | 264591 |

| | | | | | | |
|-------|-----------------------|---|--|--|-------------------|--|
| 11150 | 81325074 (2289, 2300) | Novel Protein sim. GBank gij2895095 (AF011337) - putative E1-E2 ATPase [Mus musculus] | | | ATPase-associated | 264488, 35696286, 264907, 264908, 264909, 264910, 264593, 264598, 264758, 264764, 264766, 264768, 264693, 264628, 60431850, 264564, 264566, 264587 264595 |
| 11151 | 80070874 (2301, 2302) | Novel Protein sim. GBank gj4324655[gbi AAD16978] - (AF108181) DNA polymerase III alpha subunit [Streptomyces coelicolor] Novel Protein sim. GBank gij3874275[emb CAB07311.1] - (Z02825) predicted using GeneFinder; Similarity to Yeast low affinity glucose transporter HXT4 (PS:32467); cDNA EST EMBL-C12555 comes from this gene; cDNA EST yk404c10.3 comes from this gene; cDNA EST yk404c10.5 comes from thi... | | | polymerase | |
| 11152 | 80235547 (2303, 2304) | Novel Protein sim. GBank gij3874275[emb CAB07311.1] - (Z02825) predicted using GeneFinder; Similarity to Yeast low affinity glucose transporter HXT4 (PS:32467); cDNA EST EMBL-C12555 comes from this gene; cDNA EST yk404c10.3 comes from this gene; cDNA EST yk404c10.5 comes from thi... | | | glycoprotein | 264488, 22278988, 264905, 264628, 264486 |
| 11153 | 80027783 (2305, 2306) | Novel Protein sim. GBank gij4240315[dj BAA74936.1] - (AB020720) KIAA0913 protein [Homo sapiens] | | | UNCLASSIFIED | 264910, 264555, 264557 265008 |
| 11154 | 83002995 (2307, 2308) | Novel Protein sim. GBank gij4240315[dj BAA74936.1] - (AB020720) KIAA0913 protein [Homo sapiens] | | | UNCLASSIFIED | 264603 |
| 11155 | 78411098 (2308, 2310) | Novel Protein sim. GBank gij586655[sjp P37617 ATZN_ECOLI - ZINC-TRANSPORTING ATPASE (ZNU)]-TRANSLOCATING P-TYPE ATPASE | | | UNCLASSIFIED | 264690, 264638 |
| 11156 | 57147843 (2311, 2312) | Novel Protein sim. GBank gij586655[sjp P37617 ATZN_ECOLI - ZINC-TRANSPORTING ATPASE (ZNU)]-TRANSLOCATING P-TYPE ATPASE | | | transport | 264603 |
| 11157 | 95287711 (2313, 2314) | Novel Protein sim. GBank gij418480[sjp P32139 YIHR_ECOLI - HYPOTHETICAL 34.0 KD PROTEIN IN GLNA-RBN INTERGENIC REGION | | | UNCLASSIFIED | 264906, 264907, 264758, 264768, 264769, 264689, 264638, 264566 |
| 11158 | 82454917 (2315, 2316) | Novel Protein sim. GBank gij2496481[sjp Q50724 Y09S_MYCTU - HYPOTHETICAL 87.3 KD PROTEIN CY78.27C | | | UNCLASSIFIED | 264906, 264762, 264687, 264769, 264689, 18108374, 35695855 |
| 11159 | 79188451 (2317, 2318) | Novel Protein sim. GBank gij1138408[dj BAA11490] - (D79995) similar to pig tubulin-lysozyme ligase. [Homo sapiens] | | | | 264687 |
| 11160 | 91229893 (2319, 2320) | Novel Protein sim. GBank gij1138408[dj BAA11490] - (D79995) similar to pig tubulin-lysozyme ligase. [Homo sapiens] | | | | 29331822, 29331824, 66714117, 29331826, 60433356, 265018, 265019, 83373044, 18108385, 22279000, 22279002, 264563 |
| 11161 | 74117143 (2321, 2322) | Novel Protein sim. GBank gij2443342[dj BAA22380] - (D88764) alpha 2 type I collagen [Rana catesbeiana] | | | UNCLASSIFIED | 264369 |
| 11162 | 79635357 (2323, 2324) | Novel Protein sim. GBank gij4503375[ref NP_001376.1 pDPYS - dihydropyrimidinase | | | UNCLASSIFIED | 264693 |
| 11163 | 79563186 (2325, 2326) | Novel Protein sim. GBank gij5052554[gbi AAD38607.1 AF145632] - (AF145632) BcDNA GH06032 [Drosophila melanogaster] | | | transport | 29331827, 264906 |
| 11164 | 79650828 (2327, 2328) | Novel Protein sim. GBank gij5052554[gbi AAD38607.1 AF145632] - (AF145632) BcDNA GH06032 [Drosophila melanogaster] | | | UNCLASSIFIED | 18108398, 29331827, 29331828, 29146488, 29146499, 18108354, 21906768, 29148627, 21906769, 264693, 18108382, 18108385, 264602, 264605, 264769, 18108370, 18108374, 264565 |
| 11165 | 80401868 (2329, 2330) | Novel Protein sim. GBank gij4589476[dj BAA76766.1] - (AB023139) KIAA0922 protein [Homo sapiens] | | | | 264488, 35696286, 22278989, 264259, 66714117, 60432289, 35696052, 264905, 56182435, 265006, 60433438, 264759, 21906754, 33109954, 265017, 265019, 264448, 264288, 264766, 264685, 35696423, 35695855, 264558, 18108375, 60432113 |
| 11166 | 88096456 (2331, 2332) | Novel Protein sim. GBank gij4589476[dj BAA76766.1] - (AB023139) KIAA0922 protein [Homo sapiens] | | | | |

| | | | | | |
|------|-----------------------|---|--|--------------------------------|--|
| 1167 | 78963862 (2333, 2334) | Novel Protein sim. GBank gij2580433[dbj BAA23138] - (D76414) ppGpp hydrolase [Staphylococcus aureus] | | kinase | 264488 |
| 1168 | 88094678 (2335, 2336) | | | UNCLASSIFIED | 264259, 29331827, 56182435, 60433438, 265019, 33657023, 35695855, 264568 |
| 1169 | 11805403 (2337, 2338) | | | UNCLASSIFIED | 264681 |
| 1170 | 21632244 (2339, 2340) | | | UNCLASSIFIED | 264602 |
| 1171 | 20434582 (2341, 2342) | Novel Protein sim. GBank gij2772914 (AF028249) - procollagen D [Mytilus edulis] | | UNCLASSIFIED | 264556 |
| 1172 | 78610113 (2343, 2344) | Novel Protein sim. GBank gij4757846[ref NP_004317.1 pBCL9 - B-cell CLL/lymphoma 9 | | UNCLASSIFIED | 55810764, 35696052, 264634, 264488 |
| 1173 | 80235713 (2345, 2346) | Novel Protein sim. GBank gij2564053[dbj BAA22946] - (AB007832) Bm trachealless [Bombyx mori] | | | 264508, 264906, 264907, 264909, 264591, 264632, 264638, 264639 |
| 1174 | 20283077 (2347, 2348) | Novel Protein sim. GBank gij2911027[emb CAA17520] - (AL021958) mmsA [Mycobacterium tuberculosis] | | dehydrogenase | 264600 |
| 1175 | 20711847 (2349, 2350) | Novel Protein sim. GBank gij1183333[sp P23234 DCIP_ENTCL - INDOLE-3-PYRUVATE DECARBOXYLASE (INDOLEPYRUVATE DECARBOXYLASE)] | | carboxylase | 264601 |
| 1176 | 80252645 (2351, 2352) | Novel Protein sim. GBank gij1144520 (U34956) - phosphoribosylformylglycinamide synthase [Mycobacterium tuberculosis] | | synthase | 264509, 264805, 264593, 264602, 264605 |
| 1177 | 80064647 (2353, 2354) | Novel Protein sim. GBank gij118791[sp P28643 FABG_CUPLA - 3-OXOACYL-(ACYL-CARRIER PROTEIN) REDUCTASE PRECURSOR (3-KETOACYL-ACYL CARRIER PROTEIN REDUCTASE)] | | reductase | 264605 |
| 1178 | 94128641 (2355, 2356) | Novel Protein sim. GBank gij5031697[ref NP_005594.1 pFIC1 - familial intrahepatic cholestasis 1, (progressive, Byler disease and benign recurrent)] | | ATPase-associated E1-E2 ATPase | 65274572, 18108398, 22278998, 22278999, 29331826, 264508, 264908, 264828, 33657402, 33109854, 264769, 21906765, 21908768, 21908768, 55811957, 33657023, 264629, 55811576, 35696423, 264636, 264556, 56182323, 60432113, 22279000, 22279002 |
| 1179 | 80055575 (2357, 2358) | Novel Protein sim. GBank gij2960090[emb CAA17988.1] - (AL022121) dppA [Mycobacterium tuberculosis] | | transport | 264603 |
| 1180 | 11794446 (2359, 2360) | Novel Protein sim. GBank gij2558614[emb CAA0478] - (AJ001493) dehydroquinase dehydratase [Streptomyces coelicolor] | | synthase | 264638 |
| 1181 | 17946362 (2361, 2362) | | | UNCLASSIFIED | 265017 |
| 1182 | 81494264 (2363, 2364) | Novel Protein sim. GBank gij5420387[emb CAB46678.1] - (AJ243459) proteophosphoglycan [Leishmania major] | | | 265007, 265009, 264564, 264909, 264693 |
| 1183 | 78574044 (2365, 2366) | | | | 264689, 35696423, 264638, 18108385 |
| 1184 | 52559833 (2367, 2368) | Novel Protein sim. GBank gij4091877 (AF061331) - alpha galactosidase precursor [Saccharopolyspora erythraea] | | UNCLASSIFIED | 264602 |
| 1185 | 79491185 (2369, 2370) | Novel Protein sim. GBank gij2129478[pir S51839 - chitinase (EC 3.2.1.14) precursor - beet | | glycoprotein | 263967 |

| | | | | | | | |
|------|-----------------------|---|--|--|---------------|--|--|
| 1186 | 20224012 (2371, 2372) | | | | UNCLASSIFIED | 264559 | |
| 1187 | 70248834 (2373, 2374) | | | | UNCLASSIFIED | 29331825, 265017, 18108351 | |
| 1188 | 79831387 (2375, 2376) | Novel Protein sim. GBank gi 2996039 (AF054525) - hypothetical protein [Synechococcus PCC7002] | | | UNCLASSIFIED | 264905, 264908 | |
| 1189 | 79609367 (2377, 2378) | | | | | 264692 | |
| 1190 | 78930589 (2379, 2380) | | | | UNCLASSIFIED | 265018 | |
| 1191 | 80310105 (2381, 2382) | | | | UNCLASSIFIED | 264600, 264605, 264764, 35695955, 264638, 264486 | |
| 1192 | 13521641 (2383, 2384) | | | Contains protein domain (PF01581) - FMRFamide related peptide family | | 264638 | |
| 1193 | 11103584 (2385, 2386) | | | | UNCLASSIFIED | 263978 | |
| 1194 | 78893947 (2387, 2388) | Novel Protein sim. GBank gi 854065 emb CAA58337 - (X83413) U88 [human herpesvirus 6] | | | UNCLASSIFIED | 265007, 265008 | |
| 1195 | 20445442 (2389, 2390) | Novel Protein sim. GBank gi 1790277 (AE000459) - putative oxidoreductase [Escherichia coli] | | | UNCLASSIFIED | 264605 | |
| 1196 | 13000888 (2391, 2392) | | | | | 264689 | |
| 1197 | 11392317 (2393, 2394) | Novel Protein sim. GBank gi 2497360 sp Q50715 IMDH_MYCTU - INOSINE 5'-MONOPHOSPHATE DEHYDROGENASE (IMP DEHYDROGENASE) (IMPDH) (IMPD) | | Contains protein domain (PF00571) - CBS domain | dehydrogenase | 264594 | |
| 1198 | 95290101 (2395, 2396) | | | | | 264603 | |
| 1199 | 81882011 (2397, 2398) | Novel Protein sim. GBank gi 1709525 sp P54673 P3K1_DICD1 - PHOSPHATIDYLINOSITOL 3-KINASE 1 (PI3-KINASE) (PTDINS-3-KINASE) (PI3K) | | | | 264259, 264757, 33109954, 21906768 | |
| 1200 | 9848880 (2399, 2400) | | | | | | |
| 1201 | 80503751 (2401, 2402) | Novel Protein sim. GBank gi 2499877 sp P70645 BLMH_RAT - BLEOMYCIN HYDROLASE (BLM HYDROLASE) (BMH) | | | UNCLASSIFIED | 264910 | |
| 1202 | 80082633 (2403, 2404) | Novel Protein sim. GBank gi 606342 (U18997) - ORF_0622; reading frame open far upstream of start; possible frameshift, linking to previous ORF [Escherichia coli] | | | cathepsin | 264766, 264769 | |
| 1203 | 82125373 (2405, 2406) | | | | ribosomalprot | 264600, 264558 | |
| 1204 | 80503916 (2407, 2408) | Novel Protein sim. GBank gi 2500728 sp Q59912 SECY_STRGB - PREPROTEIN TRANSLOCASE SECY SUBUNIT | | | UNCLASSIFIED | 264768, 264769, 35695917, 264910, 264760, 264908, 264907, 264629, 264908, 264909, 264766 | |
| 1205 | 80053961 (2409, 2410) | | | | | 264905, 264769, 264636 | |
| 1206 | 80241965 (2411, 2412) | | | | UNCLASSIFIED | 264566 | |
| 1207 | 79841192 (2413, 2414) | | | | UNCLASSIFIED | 264556, 264557, 264558 | |
| 1208 | 87755217 (2415, 2416) | Novel Protein sim. GBank gi 2645560 (AF027954) - Bcl-2-related ovarian killer protein [Rattus norvegicus] | | Contains protein domain (PF00452) - Apoptosis regulator proteins, Bcl-2 family | | 29331824, 264909, 265021, 18108370 | |
| | | | | | | 29331824, 264907, 264908, 265007, 264764, 264683, 264769, 264688, 264689 | |

| | | | | | |
|------|-----------------------|---|--|---------------|---|
| 1209 | 79185742 (2417, 2418) | Novel Protein sim. GBank gi1175033 sp P44398 XYLA_HAEIN - XYLOSE ISOMERASE | Contains protein domain (PF00259) - Xylose isomerase | isomerase | 264687, 264688 |
| 1210 | 56426884 (2419, 2420) | Novel Protein sim. GBank gi121095 prf S30688 - hypothetical protein o246 - Escherichia coli | | UNCLASSIFIED | 264907, 264693 |
| 1211 | 94665655 (2421, 2422) | Novel Protein sim. GBank gi3980625 emb CAB07858 - (Z93785) predicted using GeneFinder; similar to RNA recognition motif. (aka RRM, RBD, or RNP domain); cDNA EST EMBL: T01682 comes from this gene; cDNA EST EMBL: M75823 comes from this gene; cDNA EST EMBL: D27559 comes from this ge... | Contains protein domain (PF01412) - Putative GTP-ase activating protein for Arf | transferase | 264591, 264592, 264595 |
| 1212 | 79187829 (2423, 2424) | Novel Protein sim. GBank gi12143886 prf I52523 - nucleoporin p82 homolog - rat (fragment) | | | 264689, 263967 |
| 1213 | 79859633 (2425, 2426) | Novel Protein sim. GBank gi12143886 prf I52523 - nucleoporin p82 homolog - rat (fragment) | | kinase | 264909 |
| 1214 | 10144308 (2427, 2428) | Novel Protein sim. GBank gi15726285 gb AA048396.1 AF12616 - (AF126162) HERV-H LTR associating protein 2 [Homo sapiens] | | UNCLASSIFIED | 264908 |
| 1215 | 80050108 (2428, 2430) | Novel Protein sim. GBank gi2326739 emb CAB10953 - (Z98268) recN [Mycobacterium tuberculosis] | | UNCLASSIFIED | 265009, 264601, 264602, 264603, 33657109 |
| 1216 | 20438324 (2431, 2432) | Novel Protein sim. GBank gi417328 sp P33038 MURA_ENTCL - UDP-N- ACETYLGLUCOSAMINE 1- CARBOXYVINYLTRANSFERASE (ENOYLPIRUVATE TRANSFERASE) (UDP-N-ACETYLGLUCOSAMINE ENOLPYRUVYL TRANSFERASE) (EPT) | | transferase | 264604 |
| 1217 | 95011344 (2433, 2434) | Novel Protein sim. GBank gi1805460 dbj BAA08022 - (D50453) homologue of succinate semialdehyde dehydrogenase GabD of E. coli [Bacillus subtilis] | | UNCLASSIFIED | 264905, 264907, 264908, 264591, 264766, 264691, 264693, 264629, 264630, 264636, 264564 |
| 1218 | 11093680 (2435, 2436) | Novel Protein sim. GBank gi1805460 dbj BAA08022 - (D50453) homologue of succinate semialdehyde dehydrogenase GabD of E. coli [Bacillus subtilis] | | dehydrogenase | 264601 |
| 1219 | 91216252 (2437, 2438) | Novel Protein sim. GBank gi1805460 dbj BAA08022 - (D50453) homologue of succinate semialdehyde dehydrogenase GabD of E. coli [Bacillus subtilis] | | | 56181688, 29331822, 60432289, 264601, 264692, 264629 |
| 1220 | 91241524 (2439, 2440) | Novel Protein sim. GBank gi4240315 dbj BAA74936.1 - (AB020720) KIAA0913 protein [Homo sapiens] | | oncogene | 52644507, 264905, 264909, 265008, 265019, 265020, 52644150, 33657023, 264693, 33657182, 35695763, 264634, 22279000, 22279002, 264482 |
| 1221 | 83045055 (2441, 2442) | Novel Protein sim. GBank gi12143886 prf I52523 - nucleoporin p82 homolog - rat (fragment) | | UNCLASSIFIED | 264768, 265020, 264906 |
| 1222 | 20711865 (2443, 2444) | Novel Protein sim. GBank gi1730805 sp P39663 SPHR_SYN7 - ALKALINE PHOSPHATASE SYNTHESIS TRANSCRIPTIONAL REGULATORY PROTEIN SPHR | Contains protein domain (PF00486) - Transcriptional regulatory protein, C terminal | phosphatase | 264601 |
| 1223 | 11615647 (2445, 2446) | Novel Protein sim. GBank gi1172627 sp P46546 PROB_CORGL - GLUTAMATE 5- KINASE (GAMMA-GLUTAMYL KINASE) (GK) | Contains protein domain (PF01472) - PUA domain | kinase | 264593 |
| 1224 | 80432645 (2447, 2448) | Novel Protein sim. GBank gi1172627 sp P46546 PROB_CORGL - GLUTAMATE 5- KINASE (GAMMA-GLUTAMYL KINASE) (GK) | | | 264593, 264600, 264601, 264603, 264605, 264768, 18108376, 264635, 18108387 |

| | | | | | |
|------|-----------------------|--|--------------|-----------|--|
| 1225 | 80434427 (2449, 2450) | Novel Protein sim. GBank gi 2105050 emb CAB08836 - (Z95436) hypothetical protein RV3844c [Mycobacterium tuberculosis] | | | 264768 |
| 1226 | 80237518 (2451, 2452) | Novel Protein sim. GBank gi 1706768 sp P88133 FBN1_BOVIN - FIBRILLIN 1 PRECURSOR (MP340) | polymerase | | 264905, 264512, 264689 |
| 1227 | 79422138 (2453, 2454) | Novel Protein sim. GBank gi 1706768 sp P88133 FBN1_BOVIN - FIBRILLIN 1 PRECURSOR (MP340) | UNCLASSIFIED | | 264908, 264637, 264639 |
| 1228 | 79209027 (2455, 2456) | Novel Protein sim. GBank gi 1653901 dbj BAA18811 - (D90917) acriflavine resistance protein [Synecocystis sp.] | | | 264605, 264634 |
| 1229 | 94328135 (2457, 2458) | | UNCLASSIFIED | | 87168474, 265011, 87168559, 264681, |
| 1230 | 80049357 (2459, 2460) | Novel Protein sim. GBank gi 16230 sp P28598 CH60_BACSU - 60 KD CHAPERONIN (PROTEIN CPN60) (GROEL PROTEIN) | | | 264689, 264693, 65274820, 18108374 |
| 1231 | 79843141 (2461, 2462) | Novel Protein sim. GBank gi 1215733 U48718 - OphC [Agrobacterium tumefaciens] | eph | | 264909, 264605, 18108388 |
| 1232 | 79853104 (2463, 2464) | Novel Protein sim. GBank gi 16298 sp P20730 CHHC_BOMMO - CHORION CLASS HIGH-CYSTEINE HCB PROTEIN 13 PRECURSOR (HC-B.13) | UNCLASSIFIED | transport | 264908 |
| 1233 | 80255179 (2465, 2466) | Novel Protein sim. GBank gi 1215733 U48718 - OphC [Agrobacterium tumefaciens] | UNCLASSIFIED | | 264909 |
| 1234 | 79242158 (2467, 2468) | Novel Protein sim. GBank gi 29671 sp P40280 H2A_MAIZE - HISTONE H2A | UNCLASSIFIED | | 265017, 264564 |
| 1235 | 79914423 (2469, 2470) | Novel Protein sim. GBank gi 3875133 emb CAA94750 - (Z70750) similar to actin binding domain; cDNA EST EMBL:T00093 comes from this gene; cDNA EST EMBL:D34443 comes from this gene; cDNA EST EMBL:D37508 comes from this gene; cDNA EST EMBL:D84247 comes from this gene; cDNA EST EMBL... | histone | | 285008, 265010, 18108381 |
| 1236 | 81927147 (2471, 2472) | Novel Protein sim. GBank gi 3875133 emb CAA94750 - (Z70750) similar to actin binding domain; cDNA EST EMBL:T00093 comes from this gene; cDNA EST EMBL:D34443 comes from this gene; cDNA EST EMBL:D37508 comes from this gene; cDNA EST EMBL:D84247 comes from this gene; cDNA EST EMBL... | UNCLASSIFIED | | 264634, 264762 |
| 1237 | 83371782 (2473, 2474) | Novel Protein sim. GBank gi 3875133 emb CAA94750 - (Z70750) similar to actin binding domain; cDNA EST EMBL:T00093 comes from this gene; cDNA EST EMBL:D34443 comes from this gene; cDNA EST EMBL:D37508 comes from this gene; cDNA EST EMBL:D84247 comes from this gene; cDNA EST EMBL... | UNCLASSIFIED | | 265018, 55811150, 264565, 264757 |
| 1238 | 87411577 (2475, 2476) | Novel Protein sim. GBank gi 3885470 AF061443 - G protein-coupled receptor LGR4 [Rattus norvegicus] | | | 264758, 264601, 264766, 264687, 18108372, 264555, 264559 |
| 1239 | 82197449 (2477, 2478) | Novel Protein sim. GBank gi 4007990 gb AAC95339 - (AF084383) DOK protein [Mus musculus] | glycoprotein | | 264259, 28331822, 29331824, 35696052, 264508, 284908, 52844045, 52846317, 264288, 264769, 264693, 264632, 264634, 284558, 87168518, 264563 |
| 1240 | 80497259 (2479, 2480) | Novel Protein sim. GBank gi 1176192 sp P45420 YHCD_ECOLI - HYPOTHETICAL OUTER MEMBRANE USHER PROTEIN IN GLTF-NANT INTERGENIC REGION PRECURSOR | oncogene | | 284509, 264511, 264759, 264760, 264764, 264557 |
| 1241 | 80020711 (2481, 2482) | Novel Protein sim. GBank gi 121383 sp P19904 GLNA_VIBAL - GLUTAMINE SYNTHETASE (GLUTAMATE--AMMONIA LIGASE) | | | 284769 |
| 1242 | 79775890 (2483, 2484) | | UNCLASSIFIED | | 264601, 264604, 264638 |
| | | | | | 264906, 264907, 264908, 264634 |

| | | | | | |
|------|-----------------------|--|---|--------------|--|
| 1243 | 79779458 (2485, 2486) | Novel Protein sim. GBank gij3355671[emb CAA19971] - (AL031124) branched-chain amino acid aminotransferase [Streptomyces coelicolor] | | UNCLASSIFIED | 18108374, 35695917, 35695855, 265009, 264508, 264909 |
| 1244 | 10284821 (2487, 2488) | Novel Protein sim. GBank gij2970846 (AF051945) - Xin [Mus musculus] | | UNCLASSIFIED | 264691 |
| 1245 | 80437103 (2489, 2490) | Novel Protein sim. GBank gij4586338[dbj BAA76357.1] - (AB016787) cytochrome o ubiquinol oxidase B [Pseudomonas putida] | Contains protein domain (PF00115) - oxidase Cytochrome C and Quinol oxidase polypeptide I | oxidase | 264768 |
| 1246 | 80059321 (2491, 2492) | Novel Protein sim. GBank gij3581849[emb CAA20805] - (AL031541) putative phenylalanyl-HRNA synthetase beta chain [Streptomyces coelicolor] | | UNCLASSIFIED | 264604, 264636, 264557, 264564 |
| 1247 | 80084831 (2493, 2494) | Novel Protein sim. GBank gij2621684 (AE000842) - adhesion protein [Methanobacterium thermoautotrophicum] | | | 264758, 264605, 264639 |
| 1248 | 88070353 (2495, 2496) | Novel Protein sim. GBank gij1352403[sp P09467]F16P_HUMAN - FRUCTOSE-1,6-BISPHOSPHATASE (D-FRUCTOSE-1,6-BISPHOSPHATE 1-PHOSPHOHYDROLASE) (FBPASE) | Contains protein domain (PF00316) - Fructose-1,6-bisphosphatase | UNCLASSIFIED | 18108392, 264259, 29331826, 264106, 264508, 264907, 264828, 265009, 60433356, 264757, 264758, 21908754, 265010, 265011, 265018, 265019, 264760, 18108351, 18108354, 265021, 18108376, 18108377, 264630, 18108385 |
| 1249 | 80056657 (2497, 2498) | Novel Protein sim. GBank gij2791407[emb CAA16001] - (AL021184) hypothetical protein Rv1473 [Mycobacterium tuberculosis] | | transport | 264908, 265010, 264600, 264603, 264691, 18108376 |
| 1250 | 12694385 (2499, 2500) | Novel Protein sim. GBank gij112785[sp P05100]3MG1_ECOLI - DNA-3-METHYLADENINE GLYCOSIDASE I (3-METHYLADENINE-DNA GLYCOSYLASE I, CONSTITUTIVE) (TAG I) | | UNCLASSIFIED | 264689 |
| 1251 | 79850448 (2501, 2502) | | | | 264909 |
| 1252 | 79458087 (2503, 2504) | | | UNCLASSIFIED | 264683, 263976 |
| 1253 | 80050121 (2505, 2506) | Novel Protein sim. GBank gij5670178[gb AAD46616.1]AF16131 - (AF161317) NRAMP manganese transport protein Mna [Salmonella typhimurium] | | glycoprotein | 264800, 264603, 18108376 |
| 1254 | 87716767 (2507, 2508) | Novel Protein sim. GBank gij103160[pir S22126 - finger protein unkempt - fruit fly (Drosophila melanogaster)] | | UNCLASSIFIED | 35696286, 264910, 264764, 264688, 21906767, 55811957, 264692, 264556, 264639 |
| 1255 | 79189728 (2509, 2510) | | | | 264638 |
| 1256 | 87889508 (2511, 2512) | Novel Protein sim. GBank gij2985353[emb CAA04608.1] - (AJ001206) pep2 [Streptomyces coelicolor] | | UNCLASSIFIED | 60432289, 264600, 264605, 264764, 264687, 264769, 264689, 27486265, 18108374, 18108376 |
| 1257 | 80201435 (2513, 2514) | Novel Protein sim. GBank gij3193306 (AF069300) - contains similarity to Arabidopsis membrane-associated salt inducible-like protein (GB:AL021637) [Arabidopsis thaliana] | | UNCLASSIFIED | 264094, 265019 |
| 1258 | 20708150 (2515, 2516) | | | UNCLASSIFIED | 264602, 263978 |
| 1259 | 80186012 (2517, 2518) | | | UNCLASSIFIED | 264906, 264448, 264908 |
| 1260 | 80084806 (2519, 2520) | | | UNCLASSIFIED | 264634, 264639 |

| | | | | | |
|------|-----------------------|--|--|-------------------|--|
| 1261 | 67412802 (2521, 2522) | Novel Protein sim. GBank gij5689511[dbj]BAA83039.1] - (AB029010) KIAA1087 protein [Homo sapiens] | Contains protein domain (PF01699) - Sodium/calcium exchanger protein | cadherin | 29331824, 264906, 264909, 264768, 264769, 264689, 264693, 264639, 18108384, 264563, 264634 |
| 1262 | 13504589 (2523, 2524) | Novel Protein sim. GBank gij95100[pir]S21334 - hypothetical protein 4 - Agrobacterium tumefaciens | | UNCLASSIFIED | |
| 1263 | 20710997 (2525, 2526) | Novel Protein sim. GBank gij3550958 (AF004840) - CDO [Rattus norvegicus] | | struct | 264602, 264634 |
| 1264 | 80083396 (2527, 2528) | | | UNCLASSIFIED | 264563 |
| 1265 | 80253578 (2529, 2530) | | | UNCLASSIFIED | 264766, 264636, 264638, 264567 |
| 1266 | 78914604 (2531, 2532) | Novel Protein sim. GBank gij1085002[pir]S55056 - mitochondrial carrier protein DIF-1 homolog - Caenorhabditis elegans | Contains protein domain (PF00153) - Mitochondrial carrier proteins | transport | 264259, 21906754, 264389 |
| 1267 | 80558918 (2533, 2534) | Novel Protein sim. GBank gij4886445[emb]CAB43370.1] - (AL050269) hypothetical protein [Homo sapiens] | Contains protein domain (PF00583) - Acetyltransferase (GNAT) family | UNCLASSIFIED | 18108398, 22278995, 56994075, 60424289, 29331827, 264109, 264512, 265007, 265008, 265009, 264595, 33109954, 33657084, 87168559, 264600, 265018, 265019, 264369, 264688, 21906767, 265020, 52644150, 264691, 33657023, 33657349, 18108374, 264556, 18108385, 60432113, 22279002, 264486 |
| 1268 | 88178473 (2535, 2536) | Novel Protein sim. GBank gij3334791[emb]CAA19939] - (AL031107) hypothetical protein SC5A7.10c [Streptomyces coelicolor] | | UNCLASSIFIED | 264508, 264905, 264906, 264687, 264693 |
| 1269 | 79821946 (2537, 2538) | Novel Protein sim. GBank gij2851634[sp]Q50591Y0D1_MYCTU - HYPOTHETICAL 50.0 KD PROTEIN CY1A11.01 | Contains protein domain (PF01574) - IMP dehydrogenase / GMP reductase N terminus | dehydrogenase | 265010, 264601 |
| 1270 | 80031420 (2539, 2540) | Novel Protein sim. GBank gij2851634[sp]Q50591Y0D1_MYCTU - HYPOTHETICAL 50.0 KD PROTEIN CY1A11.01 | | ATPase associated | 35696052, 264908 |
| 1271 | 79840499 (2541, 2542) | Novel Protein sim. GBank gij1655665[emb]CAB03731] - (Z81368) hypothetical protein Rv2395 [Mycobacterium tuberculosis] | | UNCLASSIFIED | 264686, 264689 |
| 1272 | 79462878 (2543, 2544) | Novel Protein sim. GBank gij123726[sp]P10413[HTPG_ECOLI] - HEAT SHOCK PROTEIN HTPG (HIGH TEMPERATURE PROTEIN G) (HEAT SHOCK PROTEIN C82.5) | | UNCLASSIFIED | 264509, 264639 |
| 1273 | 80220315 (2545, 2546) | Novel Protein sim. GBank gij2129478[pir]S51939 - chitinase (EC 3.2.1.14) precursor - beet | Contains protein domain (PF00183) - Hsp90 protein | UNCLASSIFIED | 264905, 264908, 264909, 264769, 264602 |
| 1274 | 95010802 (2547, 2548) | Novel Protein sim. GBank gij1175473[sp]P44555[YAAJ_HAEIN] - HYPOTHETICAL PROTEIN H10183 | | UNCLASSIFIED | 264369 |
| 1275 | 20730763 (2549, 2550) | Novel Protein sim. GBank gij1175473[sp]P44555[YAAJ_HAEIN] - HYPOTHETICAL PROTEIN H10183 | | UNCLASSIFIED | 264556, 264603 |
| 1276 | 21148844 (2551, 2552) | Novel Protein sim. GBank gij1929513 [U64318] - ATP synthase subunit beta [Mycobacterium thermorescens] | | UNCLASSIFIED | 264605 |
| 1277 | 20438195 (2553, 2554) | Novel Protein sim. GBank gij4938504[emb]CAB43862.1] - putative protein [Arabidopsis thaliana] | | synthase | |
| 1278 | 11088385 (2555, 2556) | Novel Protein sim. GBank gij4938504[emb]CAB43862.1] - putative protein [Arabidopsis thaliana] | | struct | 263976 |

| | | | | | |
|------|-----------------------|--|---|-------------------|--|
| 1281 | 94323988 (2561, 2562) | Novel Protein sim. GBank gi 1136501 (U39546) - surface protein MCA-32 [Rattus norvegicus] | Contains protein domain (PF00047) - Immunoglobulin domain | UNCLASSIFIED | 29331825, 29331828, 264766, 83373044 |
| 1282 | 87537695 (2563, 2584) | Novel Protein sim. GBank gi 3328190 (AF074266) - proto-oncogene AF4 [Mus musculus] | | UNCLASSIFIED | 265008 |
| 1283 | 20466305 (2565, 2566) | Novel Protein sim. GBank gi 3261721 emb CAB07057 - (Z92770) hypothetical protein Rv0153c [Mycobacterium tuberculosis] | | UNCLASSIFIED | 264605 |
| 1284 | 20636325 (2567, 2568) | Novel Protein sim. GBank gi 3929022 (AF057696) - LspB [Haemophilus ducreyi] | | | 264604 |
| 1285 | 80427330 (2569, 2570) | Novel Protein sim. GBank gi 417154 sp P33126 HS82_ORYSA - HEAT SHOCK PROTEIN 82 | Contains protein domain (PF00183) - eph Hsp90 protein | | 264766, 264688, 263967 |
| 1286 | 20465254 (2571, 2572) | Novel Protein sim. GBank gi 2078004 emb CAB08451 - (Z95207) gorA [Mycobacterium tuberculosis] | reductase | | 264605, 264639 |
| 1287 | 80417530 (2573, 2574) | | | UNCLASSIFIED | 265011, 264602, 264766, 264687, 264769, 264689, 18108370, 264636, 18108385, 264563 |
| 1288 | 95338101 (2575, 2576) | Novel Protein sim. GBank gi 5353510 gb AAD2161.1 AF08891 - (AF088916) emilin precursor [Homo sapiens] | Contains protein domain (PF00386) - collagen C1q domain | | 35886052, 264107, 264508, 264509, 264905, 264906, 264907, 264908, 264909, 264510, 264511, 265007, 264512, 264910, 265009, 33657402, 264595, 264758, 265011, 265019, 264760, 18108351, 264681, 264764, 264288, 264685, 264766, 264687, 264768, 264769, 265020, 265021, 284534, 264682, 18108370, 264628, 18108374, 35896423, 264555, 264556, 264557, 264558, 18108385, 264564, 264566, 264567, 264686, 18108391, 264637 |
| 1289 | 11813647 (2577, 2578) | Novel Protein sim. GBank gi 1169995 sp P46023 GPCR_LYMT - G-PROTEIN COUPLED RECEPTOR GRL101 PRECURSOR | Contains protein domain (PF00001) - 7 transmembrane receptor (rhodopsin family) | UNCLASSIFIED | 264563 |
| 1290 | 19526027 (2579, 2580) | Novel Protein sim. GBank gi 2072674 emb CAB08305 - (Z95120) rhlE [Mycobacterium tuberculosis] | Contains protein domain (PF00271) - Helicases conserved C-terminal domain | helicase | 265007, 265008, 264769 |
| 1291 | 80470266 (2581, 2582) | Novel Protein sim. GBank gi 1835755 (U86338) - zinc finger protein Png-1 [Mus musculus] | Contains protein domain (PF01530) - Zinc finger, C2HC type | transcript factor | 264092, 264259, 29331822, 29331824, 264508, 264908, 264909, 264512, 265008, 265009, 264591, 265019, 264369, 264288, 264686, 264768, 264693, 18108374, 264632, 56182323, 264639, 83373044, 22279002, 264482, 264563 |
| 1292 | 94723316 (2583, 2584) | | | UNCLASSIFIED | 265006, 55812038, 264369, 264556 |
| 1293 | 80067536 (2585, 2586) | Novel Protein sim. GBank gi 2129173 pr F64453 - oxaloacetate decarboxylase (EC 4.1.1.3) alpha subunit - Melhanococcus jannaschii | biolindp | | 264602, 264605, 264760, 18108351, 264689, 33657023, 264559 |
| 1294 | 82125908 (2587, 2588) | Novel Protein sim. GBank gi 5441778 emb CAB46803.1 - (AL096811) putative alcohol dehydrogenase (zinc-binding) [Streptomyces coelicolor A3(2)] | dehydrogenase | | 264689 |

| | | | | | | |
|------|-----------------------|--|---|--|-------------------|--|
| 1296 | 11687804 (2591, 2592) | Novel Protein sim. GBank | | | UNCLASSIFIED | 264591, 264639 |
| 1297 | 78639300 (2593, 2594) | gi4982191gb AAD36686.1 AE00180 - (AE001805) DNA-directed DNA polymerase I [Thermotoga maritima] | Contains protein domain (PF01367) - 5'-3' exonuclease | | polymerase | 264693 |
| 1298 | 94238506 (2595, 2596) | Novel Protein sim. GBank gi1943770 (U97191) - F53F10.1 gene product [Caenorhabditis elegans] | | | struct | 18108348, 265017 |
| 1299 | 80255378 (2597, 2598) | Novel Protein sim. GBank gi3445181 (AC005498) - R31665.2 [Homo sapiens] | Contains protein domain (PF01352) - KRAB box | | transcriptfactor | 264488, 264906, 264908, 22279002, 264566 |
| 1300 | 80064667 (2599, 2600) | Novel Protein sim. GBank gi4062873dbj BAA36204.1 - (AB017138) alpha subunit of malonate decarboxylase [Pseudomonas putida] | | | UNCLASSIFIED | 264906 |
| 1302 | 95416198 (2603, 2604) | | | | | 85658542, 265020 |
| 1303 | 9684121 (2605, 2606) | | | | | 264908 |
| 1304 | 78377196 (2607, 2608) | | | | UNCLASSIFIED | 264508 |
| 1305 | 19905899 (2609, 2610) | | | | | 264566 |
| 1306 | 13069230 (2611, 2612) | Novel Protein sim. GBank gi3242273emb CAB07017 - (Z92669) hypothetical protein Rv0236c [Mycobacterium tuberculosis] | | | UNCLASSIFIED | 264636 |
| 1307 | 82201029 (2613, 2614) | Novel Protein sim. GBank gi1502421 (U59433) - 3-ketoacyl acyl carrier protein reductase [Bacillus subtilis] | Contains protein domain (PF00516) - Envelope glycoprotein GP120 | | UNCLASSIFIED | 264907, 264592, 264764 |
| 1308 | 21426814 (2615, 2616) | Novel Protein sim. GBank gi95819 pf S16298 - ferric enterobactin transport protein fepC - Escherichia coli | | | reductase | 264555 |
| 1309 | 79263011 (2617, 2618) | Novel Protein sim. GBank gi5459220 emb CAB48893.1 - (AL096837) putative iron-sulfur protein [Streptomyces coelicolor A3(2)] | | | transport | 264906, 18108354 |
| 1310 | 20466319 (2618, 2620) | | | | UNCLASSIFIED | 264605 |
| 1311 | 87613142 (2621, 2622) | | | | | 35696286, 29331827, 264908, 265008, 264764, 264766, 264686, 21908767, 21906769, 35695917, 264691, 264693, 22278995, 22278998, 22278999, 264905, 264908, 265011, 265017, 265019, 264687, 21908768, 265020, 265021, 33657023, 22279002, 264564 |
| 1312 | 88061720 (2623, 2624) | Novel Protein sim. GBank gi4455118gb AAD21084 - (AF125158) zinc finger DNA binding protein 99 [Homo sapiens] | Contains protein domain (PF00096) - Zinc finger, C2H2 type | | dna_ma_bind | 22278996, 22278999, 264259, 20281099, 29146498, 264508, 264908, 66712502, 60433356, 60433438, 265011, 265017, 264683, 264288, 21906765, 21906767, 29148827, 21906768, 35695917, 265021, 33657023, 33657109, 18108370, 18108377, 35695855, 60432113, 22279000, 264563, 18108390 |
| 1313 | 91225458 (2625, 2626) | Novel Protein sim. GBank gi4928733gb AAD34127.1 AF15189 - (AF151890) CGI-132 protein [Homo sapiens] | Contains protein domain (PF00886) - Ribosomal protein S16 | | ribosomalprot | 264693 |
| 1314 | 56926053 (2627, 2628) | Novel Protein sim. GBank gi2589223 (AF026565) - (ing finger protein [Mus musculus]) | Contains protein domain (PF00097) - Zinc finger, C3HC4 type (RING finger) | | interleukinrecept | 264691 |
| 1315 | 84357192 (2629, 2630) | | | | | |

| | | | | | |
|------|-----------------------|---|--|------------------|--|
| 1316 | 95361609 (2631, 2632) | Novel Protein sim. GBank gij5695407[dbj][BAA82987.1] - (AB028958) KIAA1035 protein [Homo sapiens] | | kinase | 56182575, 56181686, 20281171, 28331822, 29331824, 60424269, 29331825, 35696052, 52644045, 264591, 60432229, 265018, 265019, 55811150, 56181562, 21906765, 21906767, 21906768, 35695917, 60170615, 33657023, 65274620, 33657109, 35695763, 35695855, 18108387, 87168518, 60432113, 22279002, 264584 |
| 1317 | 88055167 (2633, 2634) | Novel Protein sim. GBank gij4836757[gb][AAD30541.1][AF134918] semaphorin subclass 4 member G [Mus musculus] | | UNCLASSIFIED | 264093, 264806, 264909, 264369, 264684 |
| 1318 | 95322893 (2635, 2636) | Novel Protein sim. GBank gij4680204[gb][AAD27567.1][AF11417 - (AF114171)] hypothetical protein [Sorghum bicolor] | | UNCLASSIFIED | 18108392, 18108348, 265011, 265017, 18108359, 18108362, 56182323, 18108385, 22279000 |
| 1319 | 94238546 (2637, 2638) | | | | 264908, 264909, 265006, 265008, 264592, 265019, 264766, 56181562, 18108368, 264628, 264629, 18108377, 264638 |
| 1320 | 86603567 (2639, 2640) | Novel Protein sim. GBank gij4240163[dbj][BAA74870.1] - (AB020654) KIAA0847 protein [Homo sapiens] | | UNCLASSIFIED | 35698286, 55812038, 265018, 21906768, 265020, 263978, 22279002 |
| 1321 | 86676351 (2641, 2642) | Novel Protein sim. GBank gij4886505[emb][CAB43377.1] - (AL050276) hypothetical protein [Homo sapiens] | Contains protein domain (PF00651) - BTB/POZ domain | transcriptfactor | 60432049, 29331828, 264907, 264908, 264909, 264910, 55812038, 264601, 264762, 264764, 264766, 264768, 264769, 264628, 18108374, 264634, 264635, 18108385 |
| 1322 | 87755272 (2643, 2644) | Novel Protein sim. GBank gij5262591[emb][CAB45736.1] - (AL080143) hypothetical protein [Homo sapiens] | Contains protein domain (PF00096) - Zinc finger, C2H2 type | dna_bind | 29331828, 264908, 265020, 33657023, 264693, 264404 |
| 1323 | 94845931 (2645, 2646) | Novel Protein sim. GBank gij5459516[dbj][BAA82407.1] - (AB029821) phosphatidylethanolamine N-methyltransferase [Homo sapiens] | | synthase | 65274572, 56994075, 264259, 29331822, 29331827, 264104, 56182435, 87168474, 18108351, 264288, 21906768, 21906767, 35695917, 265020, 264693, 65274781, 56182323, 18108387 |
| 1324 | 87737614 (2647, 2648) | Novel Protein sim. GBank gij5031717[ref][NP_005704.1]GPBP - goodpasture antigen-binding protein | Contains protein domain (PF01852) - START domain | | 22278996, 22278998, 29331828, 264905, 264907, 29331830, 264908, 264510, 265008, 264595, 264759, 21906754, 265018, 264288, 264768, 264769, 21906768, 265022, 18108376, 264631, 264632, 264634, 264636, 264638, 264563, 264564, 264565, 264566 |
| 1325 | 94847471 (2649, 2650) | Novel Protein sim. GBank gij3294501 (U64857) - similar to the DPTUKunitz family of inhibitors: most similar to tissue factor pathway inhibitor precursor [Caenorhabditis elegans] | Contains protein domain (PF00090) - Thrombospondin type 1 domain | protease | 35696286, 264905, 264908, 264907, 264908, 264909, 264910, 264593, 33657402, 264758, 85658542, 264760, 264768, 264769, 264691, 35696423 |
| 1326 | 87316289 (2651, 2652) | Novel Protein sim. GBank gij1397275 (U61947) - C06G3.8 gene product [Caenorhabditis elegans] | | UNCLASSIFIED | 264259, 66712502, 264682, 264683, 264635 |

| | | | | | |
|------|-----------------------|--|--|--------------|---|
| 1327 | 85322897 (2653, 2654) | Novel Protein sim. GBank gi 728321sp P39189 ALU2_HUMAN - III ALU SUBFAMILY SB WARNING ENTRY III | Contains protein domain (PF00279) - Plant lipid transfer protein family | UNCLASSIFIED | 18108398, 22278996, 22278997, 22278999, 264091, 264259, 29331822, 29331824, 29331825, 29331826, 29331827, 29331828, 284105, 264905, 56182435, 264112, 265008, 265009, 21906754, 265010, 265011, 265017, 265019, 264681, 264448, 264764, 264684, 264288, 264685, 264768, 264688, 21906767, 21906769, 29148629, 265020, 264690, 264691, 264692, 264693, 263867, 33657109, 33657182, 27486262, 33657349, 18108370, 18108374, 55810764, 35895855, 264634, 56182323, 83373044, 87168518, 60432113, 22279000, 22279002, 264553, 264567, 264488, 22278997, 29331826, 264595, 18108351, 264766, 22279002, 264482, 264567 |
| 1328 | 87753493 (2655, 2656) | | | UNCLASSIFIED | 22278996, 29331827, 264684, 264692, 33657109 |
| 1329 | 87755276 (2657, 2658) | Novel Protein sim. GBank gi 4678224 gb AA026969.1 AC00713 - (AC007135) unknown protein [Arabidopsis thaliana] | | UNCLASSIFIED | 264259, 29331825, 264512, 265019, 265021, 264555, 264558, 56526486 |
| 1330 | 87727737 (2659, 2660) | Novel Protein sim. GBank gi 437310 (L23504) - nodulin [Medicago truncatula] | | UNCLASSIFIED | 264259, 29331826, 29331827, 35696052, 29331828, 60170831, 264448, 264688, 21906765, 55811957, 265020, 33657023, 33657109, 263973, 55811576, 35696423, 35695855, 56182323 |
| 1331 | 87376764 (2661, 2662) | Novel Protein sim. GBank gi 4589586 dbj BAA76815.1 - (AB023188) KIAA0971 protein [Homo sapiens] | | UNCLASSIFIED | 65274572, 22278996, 56994075, 22278999, 60432049, 264259, 29331822, 29331826, 60432289, 29331827, 35696052, 52644045, 56182435, 264510, 21906754, 87168559, 265018, 265019, 264448, 264288, 264369, 264688, 21906765, 21906766, 21906767, 21906768, 265020, 265021, 52644150, 33657023, 56182323, 18108387, 60432113, 22279002 |
| 1332 | 84845937 (2663, 2664) | Novel Protein sim. GBank gi 5459516 dbj BAA82407.1 - (AB029821) phosphatidylethanolamine N-methyltransferase [Homo sapiens] | synthase | | 60432289, 66712502, 264591, 60433356, 60433438, 55812038, 265010, 264639, 56526486 |
| 1333 | 86098476 (2665, 2666) | Novel Protein sim. GBank gi 5689527 dbj BAA83047.1 - (AB029018) KIAA1095 protein [Homo sapiens] | | UNCLASSIFIED | 264905 |
| 1334 | 87592388 (2667, 2668) | Novel Protein sim. GBank gi 2662536 (AF036685) - Similar to protein-tyrosine phosphatase [Caenorhabditis elegans] | phosphatase | | 22278998, 22278999, 29331827, 264509, 264511, 265007, 265008, 265009, 60433438, 21906754, 87168559, 265017, 264288, 21906765, 21906767, 21906768, 21906769, 265020, 33657109, 27486264, 18108374, 264556, 264638, 264557, 60170394, 264559, 18108385, 264563 |
| 1335 | 87644798 (2669, 2670) | Novel Protein sim. GBank gi 4240285 dbj BAA74921.1 - (AB020705) KIAA0898 protein [Homo sapiens] | Contains protein domain (PF00843) - B-box zinc finger. | UNCLASSIFIED | |

| | | | | | | |
|------|-----------------------|---|---|--------------|--------------|---|
| 1336 | 87787890 (2671, 2672) | Novel Protein sim. GBank gi 465445 sp P33485 VNUA_PRVKA - PROBABLE NUCLEAR ANTIGEN | | | UNCLASSIFIED | 264509, 264905, 264512, 264764, 264693, 264635, 264637 |
| 1337 | 94312042 (2673, 2674) | Novel Protein sim. GBank gi 5689471 dbj BAA83019.1 - (AB028990) KIAA1067 protein [Homo sapiens] | | | UNCLASSIFIED | 56182575, 56994075, 22278998, 22278999, 264092, 264259, 60432289, 29331826, 264906, 264908, 264909, 264112, 265008, 285008, 60433356, 55812038, 33657084, 265011, 265017, 265018, 265019, 264682, 264448, 264883, 264369, 264688, 264689, 21806766, 21906769, 265020, 264691, 27489261, 20281089, 18108379, 55811576, 35695855, 56182323, 60432113, 22279002, 264567 |
| 1338 | 80366114 (2675, 2676) | | | | UNCLASSIFIED | 29331822, 265010, 264288, 264689, 18108370, 35695855 |
| 1339 | 80249231 (2677, 2678) | Novel Protein sim. GBank gi 1176422 (U43194) - rhophilin [Mus musculus] | | | UNCLASSIFIED | 35696052, 264909, 264688, 264556, 264558 |
| 1340 | 88316311 (2678, 2680) | | | | | 264905, 264907, 87168559, 264764 |
| 1341 | 86101485 (2681, 2682) | | | | | 264681, 264685, 264686, 264692 |
| 1342 | 80089017 (2683, 2684) | Novel Protein sim. GBank gi 5019564 emb CAB44507.1 - (AL035542) dJ994E9.5 (hs6M1-17 (novel 7 transmembrane receptor (rhodopsin family) (olfactory receptor like protein)) [Homo sapiens]) | Contains protein domain (PF00001) - im7 7 transmembrane receptor (rhodopsin family) | | | 264629 |
| 1343 | 80082862 (2685, 2686) | Novel Protein sim. GBank gi 4557543 ref NP_001384.1 pECM2 - extracellular matrix protein 2 | Contains protein domain (PF00560) Leucine Rich Repeat | struct | | 264910, 264686, 264534 |
| 1344 | 20562559 (2687, 2688) | Novel Protein sim. GBank gi 2144101 pir I55210 - Intracellular carrier - rat (fragment) | | | glycoprotein | 263978 |
| 1345 | 91225546 (2689, 2690) | Novel Protein sim. GBank gi 3881052 emb CAA19523 - (AL023043) predicted using Genefinder; similar to serine/threonine kinase; cDNA EST yk246a12.3 comes from this gene; cDNA EST yk356c10.5 comes from this gene; cDNA EST EMBL:M89047 comes from this gene; cDNA EST yk246a12.5 comes... | Contains protein domain (PF00069) - Eukaryotic protein kinase domain | kinase | | 22278998, 264907, 264681, 264685, 264689, 265020, 264693, 22279000, 22279002, 264566 |
| 1346 | 80255717 (2691, 2692) | | | | | |
| 1347 | 80417393 (2693, 2694) | Novel Protein sim. GBank gi 4504378 ref NP_003658.1 pHG38 - orphan G protein- coupled receptor HG38 | Contains protein domain (PF00560) - Leucine Rich Repeat | glycoprotein | | 264906, 264908, 264909, 265006, 264910, 265011, 265017, 264764, 264766, 264767, 264769, 264631, 264634, 264638, 264567, 264486 |
| 1348 | 87352335 (2695, 2696) | Novel Protein sim. GBank gi 3399720 dbj BAA32100 - (AB010999) peptidylarginine deiminase type IV [Rattus norvegicus] | | | UNCLASSIFIED | 264488, 264489, 264508, 264509, 264510, 264511, 264512, 264591, 264592, 264601, 264684, 264685, 264769, 264532, 264534, 264555, 264556, 264557, 264558, 22279002, 264486 |

| | | | | | |
|------|-----------------------|--|---|--------------|--|
| 1349 | 91225548 (2697, 2698) | Novel Protein sim. GBank gij2144101[prj]I55210 - Irticarbonylate carrier - rat (fragment) | | UNCLASSIFIED | 52646842, 35696286, 22278996, 22278998, 22278999, 264259, 29331822, 29331824, 35696052, 264509, 264905, 264906, 264907, 264908, 264909, 264511, 265008, 264512, 264910, 60170831, 264591, 60433438, 264757, 21906754, 265017, 265018, 264605, 264760, 264762, 264288, 264768, 264689, 21906765, 21906766, 21906767, 21906768, 55811957, 35695917, 265020, 264534, 264691, 264692, 33657023, 264693, 33657349, 18108374, 18108376, 35696423, 60170394, 22279000, 22279002, 264563, 264564 |
| 1350 | 87093136 (2689, 2700) | | | UNCLASSIFIED | 52646842, 264259, 29331825, 264908, 264511, 264604, 264288, 21906769, 265020, 33657182, 33657349, 18108374, 35695855, 264555, 264558, 18108385, 22279002, 264486 |
| 1351 | 87361327 (2701, 2702) | Novel Protein sim. GBank gij4887239[gb]AAD32246.1] - (AF084584) BAW protein [Fugu rubripes] | | UNCLASSIFIED | 264906, 264907, 264638 |
| 1352 | 80076386 (2703, 2704) | Novel Protein sim. GBank gij2144101[prj]I55210 - Irticarbonylate carrier - rat (fragment) | | UNCLASSIFIED | 284693, 263981 |
| 1353 | 95345417 (2705, 2706) | Novel Protein sim. GBank gij2144101[prj]I55210 - Irticarbonylate carrier - rat (fragment) | | | 35696286, 60424269, 264905, 264509, 264906, 264907, 264908, 264909, 264511, 264512, 264910, 264758, 264596, 55811386, 265011, 264605, 55811150, 264762, 264764, 264766, 52644229, 56181562, 35695917, 265022, 33657023, 264693, 35695783, 60431528, 264629, 263978, 35696423, 35695855, 264630, 264634, 264635, 264636, 264637, 264638, 264639, 18108385, 264563, 264564, 264566 |
| 1354 | 95350845 (2707, 2708) | Novel Protein sim. GBank gij4689108[gb]AAD27763.1[AF07703] - (AF077030) hypothetical 43.2 kDa protein [Homo sapiens] | | UNCLASSIFIED | 22278995, 22278999, 29331826, 264906, 265008, 33657402, 21906754, 265011, 87168559, 264684, 264369, 264769, 264689, 21906765, 21906768, 52644150, 33657023, 264692, 264693, 18108374, 83373044, 87168518, 22279000 |
| 1355 | 86260186 (2709, 2710) | Novel Protein sim. GBank gij1469199[dbj]BAA09487] - (D50928) The KIAA0138 gene product is novel. [Homo sapiens] | Contains protein domain (PF00076) - RNA recognition motif. (a.k.a. RRM, RBD, or RNP domain) | dna_ma_bind | 22278996, 22278997, 264259, 66714117, 264511, 21906754, 265010, 264769, 264689, 21906765, 21906768, 21906769, 284532, 27486262, 264629, 264638, 264556, 264638, 264639, 264482, 264484 |

| | | | | | | |
|------|-----------------------|---|--|--|--------------|--|
| 1356 | 95313991 (2711, 2712) | Novel Protein sim. GBank gii113865 (U40342) - ninein [Mus musculus] | | | struct | 18108387, 22278995, 22278998, 22278998, 264094, 29331828, 264905, 265006, 265007, 265008, 265010, 265017, 265018, 265019, 264764, 18108354, 264689, 21906785, 265022, 18108364, 35686423, 83373044, 18108387 |
| 1357 | 98260268 (2713, 2714) | Novel Protein sim. GBank gii897693[emb]CAA90330] - (Z50028) phosphatidylcholine transfer protein [Bos taurus] | | Contains protein domain (PF01852) - START domain | | 264259, 29331822, 29331825, 264510, 87168559, 265018, 264448, 264288, 21906785, 21806766, 21808768, 285021, 264693, 18108376 |
| 1358 | 38718455 (2715, 2716) | Novel Protein sim. GBank gii556219 (L36831) - transcription regulator [Mus musculus] | | | | 264757 |
| 1359 | 87771643 (2717, 2718) | | | | UNCLASSIFIED | 264907, 264909, 264510, 264511, 264512, 18108351, 264764, 264534, 33657023, 18108374, 264634, 264635, 264638, 264639, 18108385, 264486, 264587 |
| 1360 | 87738272 (2719, 2720) | Novel Protein sim. GBank gii2588282[emb]CAA75612] - (Y15417) acetate-CoA ligase [Coprinus cinereus] | | | synthase | 60432289, 264605 |
| 1361 | 87593527 (2721, 2722) | Novel Protein sim. GBank gii5689443[db]BAA83005.1] - (AB028976) KIAA1053 protein [Homo sapiens] | | Contains protein domain (PF00536) - SAM domain (Sterile alpha motif) | UNCLASSIFIED | 35696286, 22278997, 22278999, 264259, 29331826, 264508, 264509, 264905, 264907, 264908, 265007, 265009, 33109954, 21908754, 87168474, 265011, 264761, 264683, 264288, 264768, 264769, 264689, 21908768, 265020, 265021, 33657023, 55811576, 35696423, 264634, 60432113, 22279002, 264482, 264486 |
| 1362 | 95287961 (2723, 2724) | Novel Protein sim. GBank gii5689411[db]BAA82989.1] - (AB028960) KIAA1037 protein [Homo sapiens] | | Contains protein domain (PF00400) - WD domain, G-beta repeat | eph | 56182575, 56181686, 60432049, 264259, 29331822, 56182181, 29331827, 35696052, 29331828, 264905, 264906, 264908, 264595, 55812038, 85658542, 55811150, 264681, 264288, 264369, 56181582, 60431528, 55810764, 35696423, 60431850, 264558 |
| 1363 | 85758476 (2725, 2726) | Novel Protein sim. GBank gii1130494 (U35776) - ADP-ribosylation factor 1-directed GTPase activating protein [Rattus norvegicus] | | Contains protein domain (PF01412) - Putative GTPase activating protein for Arf | UNCLASSIFIED | 264488, 29331826, 264907, 264687, 264689, 264693 |
| 1364 | 88178488 (2727, 2728) | | | | | 60432289, 60433356, 60433438, 87168559, 264603, 18108351, 21808766, 35696423, 60432113 |
| 1365 | 83003108 (2729, 2730) | Novel Protein sim. GBank gii4589562[db]BAA76803.1] - (AB023176) KIAA0959 protein [Homo sapiens] | | | oncogene | 264766 |
| 1366 | 87003262 (2731, 2732) | Novel Protein sim. GBank gii1084944[pir]S54495 - hypothetical protein YPR021c - yeast (Saccharomyces cerevisiae) | | Contains protein domain (PF00153) - Mitochondrial carrier proteins | transport | 265007 |

| | | | | | |
|------|-----------------------|--|---|-------------------|--|
| 1367 | 87721210 (2733, 2734) | Novel Protein sim. GBank gij488408[emb](CAB43240.1) - (AL050018) hypothetical protein [Homo sapiens] | Contains protein domain (PF01342) - SAND domain | UNCLASSIFIED | 264488, 52646842, 52646365, 22278995, 56984075, 35696286, 22278996, 22278998, 22278999, 264259, 29331824, 66714117, 29331825, 60432289, 35696052, 264805, 264907, 264908, 264909, 264510, 265006, 265007, 264512, 264910, 265009, 60170831, 33657402, 55812038, 21908754, 265011, 87168559, 265017, 265019, 18108351, 264448, 264682, 264683, 264288, 264368, 264686, 264767, 264688, 21906765, 21906766, 21906768, 21906769, 55811957, 265020, 265021, 265022, 264534, 60170615, 264690, 264691, 18108362, 33657023, 33657109, 33657349, 264628, 18108370, 18108374, 18108378, 55811576, 35696423, 35695855, 264635, 264555, 264637, 264556, 52644332, 60170394, 264558, 18108381, 18108385, 56526486, 22279000, 264583, 264587 |
| 1368 | 94320078 (2735, 2736) | Novel Protein sim. GBank gij464561[sp]P35289[RB15_RAT RAS-RELATED PROTEIN RAB-15] | Contains protein domain (PF00071) - Ras family | oncogene | 284259, 29331822, 29331826, 60432289, 29331827, 35696052, 264508, 264905, 264906, 264908, 264909, 264510, 265007, 264910, 60433438, 264758, 85658542, 87168559, 264600, 264601, 264760, 264764, 264765, 264768, 52644229, 264689, 35695917, 265020, 265021, 264631, 264632, 264634, 264637, 52644332, 264558, 264639, 83373044, 264563, 264588, 264486, 264567, 265008, 60432229, 60433356, 33657084, 21906764, 21906768, 264555, 264638, 264559, 264567 |
| 1369 | 86634033 (2737, 2738) | Novel Protein sim. GBank gij2062702 (U90550) - butyrophilin [Homo sapiens] | | UNCLASSIFIED | |
| 1370 | 95316910 (2739, 2740) | Novel Protein sim. GBank gij5031823[ref]NP_005823.1pKCNM - potassium large conductance calcium-activated channel, subfamily M, beta member 2 | | potassium_channel | 22278996, 264259, 29331822, 29331824, 29331826, 29331827, 35696052, 264508, 264509, 264905, 264908, 264907, 264908, 264909, 264510, 264511, 264512, 264758, 265011, 265019, 264764, 264766, 264768, 21906767, 35695917, 18108362, 35696423, 264632, 264635, 264636, 264555, 264638, 264556, 264639, 18108385, 65274727, 264404, 264563, 264566, 264488 |
| 1371 | 95336512 (2741, 2742) | Novel Protein sim. GBank gij5032203[ref]NP_005714.1pTSPA - tetraspan 5 | Contains protein domain (PF00335) - 4 transmembrane segments integral membrane proteins | glycoprotein | 22278996, 264259, 29331822, 29331824, 29331825, 29331826, 29331827, 264905, 264509, 29331830, 66712502, 265008, 265009, 264758, 33657084, 85658542, 265010, 265018, 265019, 264762, 264448, 35695917, 33657109, 33657182, 33657349, 35695855, 264558, 22279002, 264563 |

| | | | | | | |
|------|-----------------------|---|--|-------------------|--------|--|
| 1372 | 80248517 (2743, 2744) | Novel Protein sim. GBank gij840708[dj]BAA093341 - (D50685) trans-sialidase [Trypanosoma cruzi] | | collagen | 263978 | 264769, 21906765, 21906767, 22278999, 264691, 264910, 55812038, 265010, 264681, 264684 |
| 1373 | 80499421 (2745, 2746) | | | UNCLASSIFIED | | 264907, 264510, 265007, 264512, 265008, 60432229, 264689, 65274781, 264555, 264556, 264557, 83373044, 60432113 |
| 1374 | 95087036 (2747, 2748) | Novel Protein sim. GBank gij111876[pjrl]JC1241 - beta-interferon-induced protein - rat | | interferon | | 52044507, 52845156, 52846842, 52846365, 56182575, 56181686, 22278998, 56994075, 35698286, 22278997, 22278998, 22278999, 264259, 29331822, 52645080, 29331824, 29331825, 60432289, 29331826, 29331827, 29331828, 35696052, 33656970, 264905, 264509, 264906, 264907, 264908, 29331830, 264909, 265006, 264511, 265007, 265008, 265009, 33657402, 60433358, 52646317, 33109954, 33657084, 52644298, 87168474, 87168559, 264600, 265017, 265018, 265019, 55811150, 18108351, 264448, 264764, 264286, 264369, 264766, 52644229, 21906765, 21906766, 21906767, 21906768, 21906769, 55811957, 35695917, 265020, 265021, 52644150, 33657023, 264693, 65274620, 52645129, 33657109, 27486261, 33657349, 27486265, 35695763, 18108376, 55810764, 35698423, 35695855, 264630, 264631, 264634, 264636, 264555, 264638, 18108385, 87168518, 60432113, 22279000, 22279002, 264583, 264564, 264568, 264587 |
| 1375 | 94236942 (2748, 2750) | Novel Protein sim. GBank gij5649176[gj]AAD03500.21 - (AF051155) G beta-like protein GBL [Rattus norvegicus] | Contains protein domain (PF00400) - WD domain, G-beta repeat | kinase | | 264768, 264769, 35695917, 22278997, 264691, 264259, 29331822, 264693, 35696052, 264508, 264509, 264905, 264906, 264628, 264908, 264629, 18108372, 264909, 264510, 264511, 264512, 265008, 264630, 264631, 264910, 264632, 264634, 264635, 264636, 264591, 264592, 264637, 264638, 264558, 264639, 33657402, 264595, 18108385, 56526488, 265010, 265011, 264600, 264563, 264762, 264564, 264565, 264764, 264488, 264766 |
| 1376 | 87399050 (2751, 2752) | Novel Protein sim. GBank gij138350[sp]P28968[VGLX_HSVB - GLYCOPROTEIN X PRECURSOR | | UNCLASSIFIED | | 29331824, 264591, 265019, 264686, 264768, 55811957, 264693, 22279002 |
| 1377 | 66864242 (2753, 2754) | Novel Protein sim. GBank gij1663648 (U75321) - chromaffin granule ATPase II homolog [Mus musculus] | Contains protein domain (PF00122) - E1-E2 ATPase | ATPase_associated | | |

| | | | | | |
|------|-----------------------|---|---|--------------|--|
| 1378 | 87595071 (2755, 2756) | Novel Protein sim. GBank gji4107015[db][BAA36293] - (AB001772) PEM-5 [Clona savignyi] | | | 22278994, 22278998, 264093, 264094, 264259, 29331824, 29331827, 265009, 265018, 265019, 18108351, 264369, 264288, 29148627, 55811957, 264691, 18108366, 33657109, 18108368, 264635, 263981, 18108385 |
| 1379 | 85879344 (2757, 2758) | Novel Protein sim. GBank gji3252872 (AF035620) - BRCA1-associated protein 2 [Homo sapiens] | | | 265020 |
| 1380 | 87627862 (2759, 2760) | Novel Protein sim. GBank gji483737[gb][AAD30662.1] - (AF086834) germ cell specific Y-box binding protein [Homo sapiens] | | nuc_rept | 264510, 264512, 265008, 264288, 264564 |
| 1381 | 88178656 (2761, 2762) | Novel Protein sim. GBank gji4731580[gb][AAD28508.1/AF12538 - (AF125384) L82A [Drosophila melanogaster] | | UNCLASSIFIED | 87168559, 265017, 264628, 22279002 |
| 1382 | 84847576 (2763, 2764) | | Contains protein domain (PF00583) - Acetyltransferase (GNAT) family | UNCLASSIFIED | 22278994, 22278997, 22278999, 29331822, 29331824, 29331826, 265007, 60432229, 60433356, 85858542, 265017, 265018, 264685, 264768, 21906766, 35695917, 33657023, 27486261, 27486262, 35695763, 35695855, 87168518, 22279002 |
| 1383 | 87860598 (2765, 2766) | | | UNCLASSIFIED | 18108396, 264692 |
| 1384 | 86915895 (2767, 2768) | | | UNCLASSIFIED | 264488, 264508, 264509, 264905, 264908, 264908, 264909, 264511, 264512, 264910, 264760, 18108351, 264766, 264769, 35695855, 264630, 264636, 264555, 264638, 264483, 264584, 264486 |
| 1385 | 86378788 (2769, 2770) | Novel Protein sim. GBank gji2384732 (AF015911) - NAC-1 protein [Rattus norvegicus] | | UNCLASSIFIED | 35696052, 55811386, 264688, 21906765, 265020, 33657023, 18108385 |
| 1386 | 91013049 (2771, 2772) | Novel Protein sim. GBank gji2384910 (AF022982) - contains similarity to the A-type potassium current class of channel proteins [Caenorhabditis elegans] | | Inf | 60432289, 29331828, 264906, 264907, 56182435, 265011, 264681, 60170615, 33657023, 83373044, 264566 |
| 1387 | 87797958 (2773, 2774) | Novel Protein sim. GBank gji4160304[emb][CAA10600] - (AJ132192) HS1 binding protein 3 [Mus musculus] | | UNCLASSIFIED | 264591 |
| 1388 | 95101652 (2775, 2776) | Novel Protein sim. GBank gji4895184[gb][AAD32753.1/AC007231] putative disease resistance protein [Arabidopsis thaliana] | | glycoprotein | 65274572, 22278999, 264259, 29331826, 29331827, 35696052, 264509, 264907, 264908, 264909, 265006, 265008, 60170831, 33657402, 60433438, 264596, 21906754, 87168559, 264600, 265017, 264683, 18108354, 52644229, 21906765, 21906766, 21906767, 21906768, 21906769, 265021, 264692, 33657023, 33657109, 35695855, 264558, 60170394, 83373044, 22278000 |

| | | | | | |
|------|-----------------------|---|--|--------------|--|
| 1388 | 91256016 (2777, 2778) | Novel Protein sim. GBank gij5689387[dbj BAA82977.1] - (AB028948) KIAA1025 protein [Homo sapiens] | Contains protein domain (PF00641) - Zn-finger in Ran binding protein and others. | UNCLASSIFIED | 65274572, 22278999, 264259, 29331822, 29331824, 29331825, 60432289, 29331828, 29148499, 264908, 66712502, 55812038, 265017, 265018, 265019, 18108351, 264369, 21906765, 21906766, 21906767, 21906768, 265020, 265021, 264692, 33657023, 33657349, 18108370, 18108374, 55811576, 264555, 264556, 264557, 60170394, 83373044, 22279000, 264563, 264564, 52645156, 52646365, 264259, 52645080, 29331825, 29331826, 264908, 52644045, 265009, 33657084, 52644298, 87168474, 87168559, 265017, 265018, 264760, 264682, 264288, 264686, 264687, 56181562, 52644229, 21906765, 21906769, 35695917, 265020, 265021, 80170615, 52644150, 33657023, 27486281, 27486284, 85274791, 264631, 264555, 52644332, 87168518, 22279000, 264567 |
| 1380 | 94111916 (2778, 2780) | Novel Protein sim. GBank gij3702295 (AC005783) - R33083_1 [Homo sapiens] | peptidase | | |
| 1391 | 91227345 (2781, 2782) | Novel Protein sim. GBank gij1346910[sp P28650 PUA1_MOUSE - ADENYLOSUCCINATE SYNTHETASE, MUSCLE (ISOZYME (IMP--ASPARTATE LIGASE) | Contains protein domain (PF00709) - Adenylosuccinate synthetase | | 29331826, 29331828, 29331830, 264448, 264288, 33657023, 18108365, 264555, 264556, 83373044 |
| 1392 | 94311097 (2783, 2784) | Novel Protein sim. GBank gij726286 (U22394) - mSin3A [Mus musculus] | | | 52646942, 65274572, 22278994, 22278895, 35696286, 56994075, 22278997, 22278998, 22278999, 60432049, 264259, 52645080, 29331822, 29331824, 60432289, 29331826, 29331827, 35696052, 29331828, 33656970, 264907, 52644045, 265006, 265007, 265008, 60431735, 60433356, 52646317, 55811386, 52644286, 265010, 87168559, 265017, 264604, 265018, 265019, 264448, 264288, 264369, 264766, 21906764, 21906767, 35695917, 265020, 265021, 33657109, 52645129, 27486261, 27486262, 27486265, 33657349, 35695763, 18108370, 18108374, 18108376, 55811576, 35696423, 35695855, 264636, 52644332, 18108382, 18108385, 87168518, 60432113, 22279000, 264484, 264566, 18108391 |
| 1393 | 80409472 (2785, 2786) | | Contains protein domain (PF00560) - Leucine Rich Repeat | UNCLASSIFIED | 264763, 264631 |
| 1394 | 15028819 (2787, 2788) | | | UNCLASSIFIED | 264629 |
| 1395 | 95361471 (2789, 2790) | Novel Protein sim. GBank gij2274845[dbj BAA21534] - (D88461) N-WASP [Rattus rattus] | | UNCLASSIFIED | 265009, 18108381 |

| | | | | | |
|------|-----------------------|---|--------------|---|--|
| 1396 | 95363253 (2791, 2792) | Novel Protein sim. GBank gi 2135904 pir 54810 - pHLE1F1 - human | | | 22278997, 22278999, 264259, 29331825, 60432289, 29331828, 29146498, 29146499, 264907, 264908, 29331830, 264909, 265006, 265007, 265008, 265009, 60433356, 265010, 264802, 265017, 265018, 265019, 18108354, 52644229, 18108358, 21906767, 29148627, 21906768, 21906769, 29148629, 29148784, 265021, 265022, 18108368, 18108374, 56182323, 18108385, 264563, 264567, 35696286, 264907, 66712502, 264510, 35695917, 264692, 264693, 35698423, 264259, 29331822, 29331824, 29331825, 29331827, 35696052, 33656970, 87168474, 265018, 265019, 264682, 264768, 21906767, 265020, 33657023, 27486261, 55811576, 264632, 264639, 83373044, 87168518, 22279002 |
| 1397 | 87631317 (2793, 2794) | | UNCLASSIFIED | | |
| 1398 | 91233667 (2795, 2796) | Novel Protein sim. GBank gi 5420389 emb CAB46680.1 - (AJ243460) proteophosphoglycan [Leishmania major] | | | |
| 1399 | 87631076 (2797, 2798) | Novel Protein sim. GBank gi 2496887 sp Q09232 YQ22_CAEEL - HYPOTHETICAL 32.0 KD PROTEIN C09F5.2 IN CHROMOSOME III | UNCLASSIFIED | | |
| 1400 | 95419064 (2799, 2800) | Novel Protein sim. GBank gi 283920 pir S27939 - tensin - chicken | UNCLASSIFIED | Contains protein domain (PF00017) - Src homology domain 2 | 56182575, 22278994, 22278997, 264259, 29331822, 29331825, 29331826, 29331827, 29331828, 264908, 56182435, 264112, 265009, 265011, 265017, 265018, 265019, 264760, 264762, 264765, 264288, 264685, 264687, 56181562, 264769, 21906768, 21906767, 55811957, 264691, 264692, 264628, 264629, 55811576, 264634, 264555, 264637, 264557, 264638, 18108381, 264558, 18108384, 60432113, 22279000 |
| 1401 | 91226379 (2801, 2802) | Novel Protein sim. GBank gi 3256185 emb CAA15485 - (AL008635) dJ510H16.1 [Homo sapiens] | UNCLASSIFIED | Contains protein domain (PF00790) - VHS domain | 65274572, 60432289, 264809, 264758, 264788, 21906769, 22279002 |
| 1402 | 95361475 (2803, 2804) | Novel Protein sim. GBank gi 1515427 (U57523) - nel homolog [Homo sapiens] | lgf | Contains protein domain (PF00008) - EGF-like domain | 264905, 264907, 264908, 264809, 264112, 264693, 33657109, 264634 |
| 1403 | 94147933 (2805, 2806) | Novel Protein sim. GBank gi 5262615 emb CAB45747.1 - (AL080156) hypothetical protein [Homo sapiens] | | | 65274572, 66712502, 265017, 264448, 264288, 21906765, 21906769, 264693, 55811576, 65274791, 60432113 |
| 1404 | 90935393 (2807, 2808) | | UNCLASSIFIED | | 65274572, 22278998, 29331822, 29331828, 66712502, 265008, 60433438, 265017, 264693, 18108385 |

| | | | | | |
|------|-----------------------|---|--------------|--|--|
| 1405 | 95095068 (2808, 2810) | Novel Protein sim. GBank gij854065[emb]CAA58337] - (X83413) U88 [Human herpesvirus 6] | | | 264488, 56994075, 35696286, 29331822, 29331824, 29331826, 29331828, 35696052, 264508, 264908, 264907, 264908, 264510, 264511, 264910, 33657402, 264594, 264758, 264600, 264604, 264762, 18108351, 264764, 33657023, 33657109, 264628, 264634, 83373044, 22279002, 264563, 264482, 284486, 264567 |
| 1408 | 87612369 (2811, 2812) | Novel Protein sim. GBank gij624076[g]AAC96425.1] - (U42580) contains Pro-rich Px motifs: SPKPP (20X), PEPPA (9X); similar to soybean pro-rich cell wall protein, corresponds to Swiss-Prot Accession Number P13993 [Paramedum bursaria Chlorella virus 1] | collagen | | 284907, 264605 |
| 1407 | 94129872 (2813, 2814) | Novel Protein sim. GBank gij2827886 (AF015037) - endonucleopeptidase A related protein; EOPA related protein [Oryctolagus cuniculus] | UNCLASSIFIED | | 35696286, 22278999, 264094, 264259, 68714117, 29331826, 29331827, 29331828, 29146498, 264107, 264908, 265006, 265008, 264910, 60433438, 285011, 285017, 18108351, 264448, 264288, 264686, 21906765, 21908769, 264692, 33657109, 18108370, 264628, 263972, 18108374, 35696423, 55811576, 264631, 264557, 264558, 83373044, 18108385, 87168518, 60432113, 22279002 |
| 1408 | 95361477 (2815, 2816) | Novel Protein sim. GBank gij2564953 (AF030001) - unknown [Mus musculus] | oncogene | Contains protein domain (PF000008) - EGF-like domain | 264488, 264489, 35696286, 264109, 264508, 264905, 264509, 264906, 264907, 264908, 264909, 265008, 265009, 264910, 33657402, 264757, 264758, 285011, 264601, 265017, 264760, 264762, 264683, 264685, 264766, 264687, 264689, 21808767, 285021, 264690, 264691, 33657023, 264692, 264693, 33657109, 264628, 264629, 35696423, 35695855, 264631, 264632, 264634, 264635, 264555, 264636, 264637, 264638, 56182323, 264639, 264563, 264564, 264565, 264566, 264567 |
| 1409 | 66644385 (2817, 2818) | Novel Protein sim. GBank gij2662165[dbj]BAA23714] - (AB007902) HH0712 cDNA clone for KIAA0442 has a 574-bp insertion at position 1474 of the sequence of KIAA0442. [Homo sapiens] | | | 264693 |
| 1410 | 86612587 (2819, 2820) | Novel Protein sim. GBank gij2493780[sp]Q60994[ACR3] MOUSE - 30 KD ADIPOCYTE COMPLEMENT-RELATED PROTEIN PRECURSOR (ACRP30) (ADIPOCYTE SPECIFIC PROTEIN ADIPOQ) | complement | Contains protein domain (PF003386) - C1q domain | 29331826, 264112, 264512, 265009, 265010, 264601, 264688, 264769, 21906767, 263974, 264631, 264566 |

| | | | | | |
|------|-----------------------|--|---|----------------|---|
| 1411 | 87818841 (2821, 2822) | Novel Protein sim. GBank gij3123155lsp91343YJM3M_CAEEL - HYPOTHETICAL 49.0 KD TRP-ASP REPEATS CONTAINING PROTEIN F55F8.5 IN CHROMOSOME I | Contains protein domain (PF00400) - WD domain, G-beta repeat | kinasereceptor | 22278995, 22278997, 22278999, 29331822, 60432289, 29331828, 284907, 265017, 265018, 264682, 21906767, 21906768, 21906769, 265020, 264690, 264691, 33657023, 33657109, 27486284, 264628, 263972, 264634, 284558, 18108385 264757 |
| 1412 | 84390918 (2823, 2824) | | | UNCLASSIFIED | |
| 1413 | 95416559 (2825, 2826) | Novel Protein sim. GBank gij3879121emb CAA94370 - (Z70310) predicted using GeneFinder. Similarity to Mouse ankyrin (PIR Acc. No. S3771); cDNA EST EMBL: T01923 comes from this gene; cDNA EST EMBL: D32335 comes from this gene; cDNA EST EMBL: D32723 comes from this gene; cDNA EST E... | Contains protein domain (PF00023) - Ank repeat | homeobox | 56994075, 29331822, 35695052, 29331828, 29331830, 284909, 52644045, 264510, 52644296, 85658542, 87168474, 265017, 265018, 264681, 264687, 21906768, 35695917, 265020, 52644150, 264692, 263967, 27486284, 35695763, 264639, 18108387, 264566 |
| 1414 | 94675860 (2827, 2828) | Novel Protein sim. GBank gij3252981 (AF068921) - Ras- binding protein SUR-8 [Mus musculus] | Contains protein domain (PF00560) - Leucine Rich Repeat | UNCLASSIFIED | 264682, 264683, 265022, 264636 |
| 1415 | 94326948 (2829, 2830) | Novel Protein sim. GBank gij1871187 (U90439) - unknown protein [Arabidopsis thaliana] | | | 52646365, 56182575, 22278994, 22278995, 56994075, 22278996, 22278997, 22278998, 22278999, 264259, 29331822, 29331824, 29331825, 29331826, 29331827, 29331828, 29146498, 66712502, 29331830, 52644045, 264113, 264511, 33657402, 264757, 21906754, 55811388, 265017, 265018, 265019, 264761, 264683, 264369, 264289, 264686, 264689, 21906768, 21906767, 29148627, 21906769, 55811957, 265020, 265021, 264690, 33657023, 65274620, 52645129, 27486282, 27486284, 60431528, 264629, 35695855, 56182323, 264559, 60432113, 264404, 22279002, 264482 |

| | | | | | |
|------|-----------------------|--|--|-------------------|---|
| 1416 | 84325977 (2831, 2832) | Novel Protein sim. GBank gi 5106557 gb AAD39749.1 AF12305 - (AF123052) MLL sepiin-like fusion protein [Homo sapiens] | Contains protein domain (PF00735) - Cell division protein | struct | 18108392, 18108394, 18108397, 18108398, 22278995, 22278996, 35698286, 22278997, 22278998, 22278999, 29331822, 35698052, 29331828, 29146498, 264905, 264907, 264908, 264828, 264909, 264113, 265006, 265007, 265008, 265009, 60170831, 264595, 18108348, 21906754, 87168474, 265010, 265011, 87168559, 265017, 265018, 265019, 264762, 18108351, 264681, 264763, 264682, 264683, 264766, 52644229, 264688, 264689, 21906765, 21906766, 21906767, 29148627, 21906768, 55811957, 29148629, 265020, 52644150, 18108361, 33657023, 18108362, 18108368, 264628, 18108370, 264629, 18108374, 18108379, 55811576, 65274791, 264634, 264636, 56182323, 18108381, 60170394, 18108385, 56526486, 87168518, 22279000 |
| 1417 | 87826663 (2833, 2834) | Novel Protein sim. GBank gi 4958935 dbj BAA78095.1 - (AB027570) suppressor of potassium transport defect 3 (Rattus norvegicus) | | ATPase_associated | 264107, 264448 |
| 1418 | 87594276 (2835, 2836) | | | UNCLASSIFIED | 264259, 264808, 265010, 52644229, 21906764, 21906768, 264690, 264639, 18108388 |
| 1419 | 87757168 (2837, 2838) | Novel Protein sim. GBank gi 2072294 (U95097) - mitotic phosphoprotein 43 [Xenopus laevis] | | struct | 264259, 60432289, 265006, 87168474, 264288 |
| 1420 | 87298628 (2839, 2840) | Novel Protein sim. GBank | | ATPase_associated | 29331824, 265007, 264563 |
| 1421 | 94746986 (2841, 2842) | Novel Protein sim. gb 5174421 ref NP_008023.1 pCPNE - copine VI (neuronal) (Z69635) Similarity to Yeast uridine kinase (SW:URK1_YEAST); cDNA EST EMBL:Z14695 comes from this gene; cDNA EST CEMSE17F comes from this gene; cDNA EST EMBL:D67355 comes from this gene; cDNA EST yk209h1.5 comes from this gen... | | kinase | 18108358, 18108396, 18108397, 21906766, 18108398, 21906767, 56182575, 21906768, 21906769, 56181686, 55811957, 35695917, 35696286, 22278996, 22278997, 22278998, 22278999, 265021, 265022, 60170615, 264259, 33657023, 29331822, 56182181, 29331824, 66714117, 29331825, 33657109, 29331826, 27486261, 29331828, 35696052, 33657349, 264905, 264509, 20281149, 18108370, 264907, 60431528, 68712502, 263972, 55811576, 35696423, 35695855, 264312, 285007, 60431850, 60432229, 60431735, 56182323, 264558, 60170394, 83373044, 55812038, 264758, 18108385, 21906754, 55811386, 87168518, 87168559, 60432113, 265017, 265018, 265019, 22279002, 55811150, 264563, 264682, 264763, 264448, 264566, 264486, 18108391 |

| | | | | | |
|------|-----------------------|--|---|--------------|--|
| 1422 | 88178777 (2843, 2844) | Novel Protein sim. GBank gil4505939[ref]NP_000928.1lpPOLR - polymerase (RNA) II (DNA directed) polypeptide A (220kD) | | mapolymerase | 56984075, 35696286, 87168559, 55811957, 55811576, 264555, 264557, 87168518 |
| 1423 | 86997762 (2845, 2846) | | | UNCLASSIFIED | 264686, 264489, 264692, 264594, 264603, 265018, 264908 |
| 1424 | 95201610 (2847, 2848) | Novel Protein sim. GBank gil437181 (U02289) - GTPase- activating protein [Caenorhabditis elegans] | Contains protein domain (PF00620) - struct RhoGAP domain | | 29331822, 29331825, 29331827, 29146498, 264905, 264908, 264909, 265007, 264910, 265009, 33109954, 265010, 87168559, 265019, 264766, 264687, 21906765, 21906766, 21906767, 21906768, 29148627, 55811957, 29148629, 265021, 264691, 264692, 56526486, 22279002, 264583 |
| 1425 | 21662314 (2849, 2850) | Novel Protein sim. GBank gil100798[pir]S14959 - proline- rich protein - wheat | | UNCLASSIFIED | 265007, 264558 |
| 1426 | 94322115 (2851, 2852) | Novel Protein sim. GBank gil2078441 (U56964) - weak similarity to S. cerevisiae intracellular protein transport protein US1 (SP:P25386) [Caenorhabditis elegans] | | UNCLASSIFIED | 264488, 60424179, 35696286, 22278997, 22278998, 22278999, 264259, 60432049, 29331822, 29331825, 29331827, 56182435, 264910, 60433356, 60433438, 21906754, 265018, 264288, 21906765, 21906766, 21906767, 21906769, 265020, 265022, 33657109, 18108370, 18108376, 264558, 83373044, 18108385, 56526486, 22279002, 264482 |
| 1427 | 91227510 (2853, 2854) | Novel Protein sim. GBank gil5816074[gb]AAD45616.1[AF08194] prolute- derived STE20-like kinase PSK [Homo sapiens] | Contains protein domain (PF00069) - Eukaryotic protein kinase domain | kinase | 22278994, 56994075, 22278997, 29331828, 29331830, 264828, 265006, 265007, 265008, 265009, 264683, 264288, 18108354, 21906765, 21906768, 29148629, 33657023, 18108374, 35695855, 83373044, 22279002, 264584 |
| 1428 | 94323008 (2855, 2856) | Novel Protein sim. GBank gil138350[sp]P28968[VLX_HSVB - GLYCOPROTEIN X PRECURSOR | | glycoprotein | 56181686, 264259, 264907, 265007, 265009, 264595, 265010, 264686, 65274620, 264629, 65274791, 22279002, 264566 |
| 1429 | 87888689 (2857, 2858) | | | UNCLASSIFIED | 264112, 264595, 265017, 265019, 21906765, 263977, 264555 |
| 1430 | 94735021 (2859, 2860) | Novel Protein sim. GBank gil1181619[dbj]BAA11565] - (D82364) a variant of TSC-22 [Gallus gallus] | | | 264094, 29331824, 264591, 264593, 265018, 264681, 21906765, 21906767, 65274620, 55811576, 264639, 87168518, 22279002 |
| 1431 | 80429081 (2861, 2862) | Novel Protein sim. GBank gil5420389[emb]CAB46860.1] - (AJ243460) proteophosphoglycan [Leishmania major] | | UNCLASSIFIED | 264768, 265020, 264691, 264693, 264638 |
| 1432 | 87463004 (2863, 2864) | Novel Protein sim. GBank gil414797 (L18966) - pyruvate dehydrogenase phosphatase [Bos taurus] | | phosphatase | 18108394, 29146498, 265007, 60433438, 264763, 29148629, 263869 |
| 1433 | 87605403 (2865, 2866) | Novel Protein sim. GBank gil2460316 (AF022147) - uterus- ovary specific putative transmembrane protein [Rattus norvegicus] | Contains protein domain (PF00100) - Zona pellucida-like domain | UNCLASSIFIED | 264259, 264510, 264591, 264603, 264565 |
| 1434 | 85713730 (2867, 2868) | | | UNCLASSIFIED | 264682, 264691 |

| | | | | | |
|------|-----------------------|---|---|-------------------|--|
| 1435 | 94708213 (2868, 2870) | Novel Protein sim. GBank gij3970850[dbj BAA34789.1] - (AB015330) HRIHFB2007 [Homo sapiens] | | transcript factor | 22278997, 22278998, 22278999, 60432049, 264259, 29331822, 29331824, 66714117, 29331826, 33656970, 264508, 264905, 66712502, 29331830, 264909, 265007, 265008, 264910, 265009, 60433356, 60433438, 264598, 21906754, 265010, 265017, 265018, 265019, 18108351, 264762, 264448, 264288, 264769, 21906767, 21906768, 21906769, 265020, 265021, 265022, 264690, 264691, 33657109, 264628, 18108374, 18108376, 55811576, 264636, 60170394, 56182323, 264559, 83373044, 87168518, 60432113, 22279000, 22279002, 264563, 264482, 264565, 263978, 264557, 264559 |
| 1436 | 86635024 (2871, 2872) | Novel Protein sim. GBank gij3183977[emb CAA39515] - (X56044) protein HIFC [Mus musculus] | | UNCLASSIFIED | |
| 1437 | 87631082 (2873, 2874) | Novel Protein sim. GBank gij2496887[sp Q09232 YQ22_CAEEL - HYPOTHETICAL 32.0 KD PROTEIN C09F5.2 IN CHROMOSOME III | | UNCLASSIFIED | 22278997, 66714117, 29331826, 264907, 56182435, 265009, 18108351, 264692, 264693 |
| 1438 | 85544280 (2875, 2876) | Novel Protein sim. GBank gij1905906 (AD000092) - hypothetical human serine-threonine protein kinase R31240_1 [Homo sapiens] | Contains protein domain (PF00595) - PDZ domain (Also known as DHR or GLGF). | kinase | 264488, 264508, 264908, 264909, 264757, 264600, 264601, 264605, 264768, 264769, 264690, 35696423, 264558, 264563, 264566 |
| 1439 | 91231884 (2877, 2878) | Novel Protein sim. GBank gij3876299[emb CAA94892] - (Z71180) similar to BPTIKUNITZ inhibitor domain; cDNA EST EMBL:D68293 comes from this gene; cDNA EST yk448h4.5 comes from this gene; cDNA EST yk249e6.5 comes from this gene; cDNA EST yk448h4.3 comes from this gene [Caenorhabditis...] | Contains protein domain (PF00450) - Serine carboxypeptidase | cathepsin | 264489, 18108394, 65274572, 56182575, 22278994, 22278996, 35696286, 22278997, 22278998, 22278999, 264259, 60432049, 29331822, 29331824, 60432289, 29331826, 29331827, 29331828, 35696052, 33856970, 264107, 264508, 264509, 264907, 66712502, 29331830, 56182435, 264511, 265006, 265007, 265009, 60432229, 60433438, 264595, 55812038, 55811386, 265011, 265017, 265018, 265019, 18108351, 264448, 18108354, 264288, 18108355, 264767, 21906765, 21906766, 21906767, 21906768, 21906769, 55811957, 35695917, 265020, 265021, 265022, 33657109, 18108370, 18108374, 55810764, 35695855, 264634, 264638, 56182323, 83373044, 18108387, 87168518, 60432113, 22279000, 264486 |
| 1440 | 87423643 (2878, 2880) | Novel Protein sim. GBank gij2662165[dbj BAA23714] - (AB007902) HH0712 cDNA clone for KIAA0442 has a 574-bp insertion at position 1474 of the sequence of KIAA0442. [Homo sapiens] | | UNCLASSIFIED | 264887, 264259, 264906, 264907, 264908, 264909, 265008, 265010, 265017, 265018, 265019, 18108351, 264369, 265020, 33657023, 33657109, 60431528, 55811576, 264635 |

| | | | | | |
|------|-----------------------|---|---|------------------|---|
| 1441 | 95317662 (2881, 2882) | Novel Protein sim. GBank gj14893956[embjCAB11123.2] - (Z98551) predicted using hexExon; MAL3P6.28 (PFC0845c). Hypothetical protein, len: 167 aa; Similarity to model organism hypothetical proteins (C.elegans, D.melanogaster, S.cerevisiae & S.pombe). C.elegans protein ZK287.5 (TR)... | Contains protein domain (PF00646) - F-box domain. | helicase | 18108392, 264488, 263994, 264488, 56182575, 22278994, 22278995, 56994075, 35696286, 22278997, 22278999, 60432049, 264259, 29331822, 29331824, 56182181, 66714117, 29331826, 29331827, 29331828, 35696052, 29146498, 264508, 264905, 264509, 264906, 264907, 264908, 264909, 56182435, 264510, 264511, 264512, 264910, 264592, 264593, 33657402, 60433438, 264595, 264758, 21906754, 85658542, 87168474, 265010, 87168559, 264600, 264602, 265017, 264604, 265018, 264605, 265019, 264760, 264761, 264762, 264881, 264448, 264764, 264683, 264288, 264766, 264768, 264769, 52644229, 264689, 21908765, 21906766, 21906787, 21906768, 21906769, 55811957, 35695917, 265020, 265021, 601770815, 52644150, 264691, 264692, 33657023, 65274620, 33657109, 35695763, 264628, 18108370, 264629, 18108374, 55811576, 35696423, 65274791, 35695855, 264631, 264634, 264635, 264636, 264637, 264638, 56182323, 264558, 60170394, 264639, 264559, 83373044, 18108385, 18108388, 56526486, 22279000, 22279002, 264563, 264483, 264564, 264566, 264567, 264486 |
| 1442 | 83367491 (2883, 2884) | Novel Protein sim. GBank gj15103027[dbjBAA78765.1] - (AB023419) mSox7 [Mus musculus] | | transcriptfactor | 264906, 265007, 264693, 264558 |
| 1443 | 87108935 (2885, 2886) | Novel Protein sim. GBank gj14887229[gbjAAD32244.1]AF15075 - (AF150755) microtubule-actin crosslinking factor. [Mus musculus] | Contains protein domain (PF00435) - Spectrin repeat | struct | 52645080, 264691, 264628, 264555 |
| 1444 | 87620478 (2887, 2888) | Novel Protein sim. GBank gj13874447[embjCAB02772] - (Z81039) predicted using Genefinder; cDNA EST EMBL: T01209 comes from this gene; cDNA EST yk278a11.3 comes from this gene; cDNA EST yk278a11.5 comes from this gene; cDNA EST yk308a9.3 comes from this gene; cDNA EST yk308a9.5 com... | UNCLASSIFIED | UNCLASSIFIED | 264259, 29331822, 29331824, 66714117, 29331828, 264288, 35695917, 33657023, 264635, 60170394 |
| 1445 | 94990470 (2889, 2890) | Novel Protein sim. GBank gj12859886[embjCAA11022] - (AJ222968) L-perilaxin [Mus musculus] | | UNCLASSIFIED | 264369 |
| 1446 | 85079268 (2891, 2892) | Novel Protein sim. GBank | | | 264369 |
| 1447 | 86945392 (2893, 2894) | gj15081610[gbjAAD39494.1]AF13544 - (AF135440) huntington yeast partner C [Mus musculus] | Contains protein domain (PF01846) - FF domain | | 18108396, 35696286, 22278997, 66714117, 29331828, 265009, 264758, 265018, 264288, 21906766, 21906767, 264692, 264634, 264566 |
| 1448 | 94990477 (2895, 2896) | Novel Protein sim. GBank gj13980411 (AC004561) - putative proline-rich protein [Arabidopsis thaliana] | Contains protein domain (PF00439) - Bromodomain | | 29331827, 264509, 264908, 265008, 264595, 18108357, 18108385, 264566, 264486 |

| | | | | | | |
|------|-----------------------|---|--|--|--------------|--|
| 1449 | 87860859 (2897, 2898) | | | | UNCLASSIFIED | 66714117, 264908, 264908, 264591, 264601, 264764, 264832 |
| 1450 | 87458686 (2899, 2900) | Novel Protein sim. GBank gij1707074 (U80450) - M01E11.2 [Caenorhabditis elegans] | | | UNCLASSIFIED | 35898286, 35898052, 265008, 265009, 60170831, 33109954, 264683, 264689, 35898423, 35898555, 56526488 |
| 1451 | 87797970 (2801, 2802) | Novel Protein sim. GBank gij4160304[emb]CAA106001 - (AJ132192) HS1 binding protein 3 [Mus musculus] | | | UNCLASSIFIED | 29331826, 264683, 264693, 263978, 264630 |
| 1452 | 85682899 (2903, 2904) | Novel Protein sim. GBank gij2832906[dbj]BAA24608.11 - (D89340) dipeptidyl peptidase III [Rattus norvegicus] | | | peptidase | 264681, 33657023, 264629 |
| 1453 | 86130434 (2905, 2906) | Novel Protein sim. GBank gij728831[sp]P39188[ALU1_HUMAN - IIII] ALU SUBFAMILY J WARNING ENTRY IIII | | | kinase | 264510, 264768 |
| 1454 | 11204698 (2907, 2908) | | | | | 264556 |
| 1455 | 87797896 (2809, 2910) | | | | UNCLASSIFIED | 29331822, 66714117, 29331825, 264905, 29331830, 265006, 265008, 265009, 265011, 265019, 18108351, 21906768, 33657109, 18108376, 264632, 56182323, 87168518 |
| 1456 | 86320218 (2911, 2912) | Novel Protein sim. GBank gij728230[sp]P41004[CUT3_SCHPO - CHROMOSOME SEGREGATION PROTEIN CUT3] | | | transport | 22278995, 22278996, 22278997, 22278998, 22278999, 29331827, 264107, 265017, 21906765, 21906766, 21906767, 21906769, 29148629, 18108370, 22278000 |
| 1457 | 80076900 (2913, 2914) | | | | UNCLASSIFIED | 264107, 264566 |
| 1458 | 87800460 (2915, 2916) | Novel Protein sim. GBank gij2246532 (U93872) - ORF 73, contains large complex repeat CR 73 [Kaposi's sarcoma-associated herpesvirus] | | | transport | 56182575, 22278999, 60432049, 264259, 29331826, 29331827, 29331828, 264102, 264107, 264110, 265009, 60432229, 265019, 265020, 263972, 263976, 264635, 22279002, 264566 |
| 1459 | 85360920 (2917, 2918) | Novel Protein sim. GBank gij5524667[gb]AAD44333.1[AF159356] Munc13-4 protein [Rattus norvegicus] | Contains protein domain (PF00168) - C2 domain | | kinase | 22278997, 264259, 29331824, 29331826, 29331827, 29331828, 265017, 265018, 264760, 264682, 264448, 264288, 264766, 265021, 264692, 33657023, 33657109, 35898555, 264566 |
| 1460 | 95354602 (2919, 2920) | | | | UNCLASSIFIED | 29331822, 264591, 55811957, 264691, 264693, 65274620 |
| 1461 | 94741513 (2921, 2922) | Novel Protein sim. GBank gij1707274 (U80931) - strong similarity to class-III of pyridoxal-phosphate-dependent aminotransferases [Caenorhabditis elegans] | Contains protein domain (PF00202) - Aminotransferases class-III pyridoxal phosphate | | gaba | 22278997, 29331822, 35898052, 265009, 264758, 265017, 265018, 265019, 264760, 264369, 264687, 21906765, 21906768, 265022, 33657109, 27486281, 264555, 83373044 |
| 1462 | 87732018 (2923, 2924) | | | | UNCLASSIFIED | 264555, 264556 |
| 1463 | 88090605 (2925, 2926) | Novel Protein sim. GBank gij1770466[emb]CAA669121 - (X98259) M-phase phosphoprotein 8 [Homo sapiens] | Contains protein domain (PF00385) - 'chromo' (CHR) domain Organization (Modifier) domain | | struct | 60432049, 264259, 29146499, 264906, 264907, 264512, 265017, 264763, 264766, 18108370, 18108374, 264636, 18108385, 18108388 |

| | | | | | |
|------|-----------------------|--|--|--------------|--|
| 1464 | 87620482 (2827, 2828) | Novel Protein sim. GBank gij3874447[emb]CAB02772] - (Z81039) predicted using Genelinder; cDNA EST EMBL: T01209 comes from this gene; cDNA EST yk278a11.3 comes from this gene; cDNA EST yk308a9.3 comes from this gene; cDNA EST yk308a9.5 comes from this gene; cDNA EST yk308a9.5 comes from this gene; cDNA EST yk308a9.5 comes from this gene... | | UNCLASSIFIED | 264569, 22278995, 22278996, 22278997, 22278998, 29331822, 29331824, 29331825, 35696052, 20281100, 264905, 29331830, 264909, 265007, 33657402, 21908754, 265017, 265018, 264682, 264684, 264369, 264288, 264766, 21908765, 21908766, 21908767, 21908769, 35695917, 264691, 33657023, 264692, 35696423, 35695855, 264630, 264631, 264639, 264565 |
| 1465 | 87425192 (2829, 2830) | Novel Protein sim. GBank gij4589588[jb]BAA76821.1] - (AB023194) KIAA0977 protein [Homo sapiens] | | glucoamylase | 264488, 22278994, 56994075, 60432049, 264259, 56182181, 60432289, 29331827, 52844045, 264511, 265007, 265008, 264596, 55812038, 55811386, 264600, 264602, 265017, 265018, 264604, 265019, 18108351, 18108354, 56181562, 21908769, 265021, 33657023, 33657182, 55811576, 264557, 18108382, 60432113 |
| 1466 | 87606227 (2831, 2832) | Novel Protein sim. GBank gij2246532 (U93872) - ORF 73, contains large complex repeat CR 73 [Kaposi's sarcoma-associated herpesvirus] | | UNCLASSIFIED | 264512, 265017, 264689, 264558 |
| 1467 | 87614326 (2833, 2834) | Novel Protein sim. GBank gij4507241[ref]NP_003137.1pSSRP - structure specific recognition protein 1 | | struct | 264683, 264636 |
| 1468 | 95342862 (2835, 2836) | Novel Protein sim. GBank gij4507241[ref]NP_003137.1pSSRP - structure specific recognition protein 1 | | | 22278998, 264758, 265018, 265019, 21908769, 265020, 33657109, 22279002 |
| 1469 | 79236174 (2837, 2838) | Novel Protein sim. GBank gij1906586 (U81788) - kinesin-73 [Drosophila melanogaster] | | struct | 18108394, 18108397, 18108398, 35696052, 29146499, 265007, 265008, 285009, 265010, 265011, 18108354, 18108365, 18108368, 18108374, 18108381, 18108382, 18108384, 18108388 |
| 1470 | 94990482 (2839, 2840) | Novel Protein sim. GBank gij5649170[jb]AAD43131.2[AF159092] - (AF159092) sylat709613 protein [Homo sapiens] | | UNCLASSIFIED | 18108394, 18108398, 56182575, 264259, 29331822, 29331824, 29331825, 60432289, 264907, 264909, 265007, 264910, 265009, 264591, 60432229, 60433356, 264595, 60433438, 264758, 33109954, 265010, 265011, 265018, 264760, 264448, 264764, 264288, 264369, 18108357, 264769, 18108358, 21908767, 21908769, 55811957, 265021, 18108361, 264691, 18108362, 18108365, 18108368, 264628, 18108379, 264637, 264557, 18108381, 56182323, 18108382, 83373044, 18108384, 18108388, 87168518, 60432113, 264404, 22279002, 264482, 264567, 264487 |
| 1471 | 87826842 (2841, 2842) | Novel Protein sim. GBank gij3876148[emb]CAB01750] - (Z78542) similar to Mitochondrial carrier proteins; cDNA EST EMBL: T01651 comes from this gene [Caenorhabditis elegans] | Contains protein domain (PF00153) - transport Mitochondrial carrier proteins | transport | 29331822, 29331824, 29331825, 264828, 264603, 264689, 264693, 18108374, 55811576 |

| | | | | | |
|------|-----------------------|--|--|--------------|--|
| 1472 | 87756816 (2943, 2944) | Novel Protein sim. GBank gi 4680707 gb AAD27743.1 AF13298 - (AF132988) CGI-34 protein [Homo sapiens] | | UNCLASSIFIED | 264905, 18108351, 21906765, 264486 |
| 1473 | 87781609 (2945, 2946) | Novel Protein sim. GBank gi 3688780 (AF042180) - testis-specific Y-encoded-like protein [Mus musculus] | Contains protein domain (PF00956) - MHC Nucleosome assembly protein (NAP) | | 18108394, 22278995, 56994075, 22278999, 29331822, 29331824, 68714117, 29331825, 29331826, 35696052, 264906, 264907, 56182435, 265007, 264758, 265018, 265019, 264760, 264764, 264288, 264685, 264686, 264768, 21906769, 55811957, 265021, 264691, 264693, 264629, 55811576, 264634, 264638, 58182323, 22279002, 264566, 264486 |
| 1474 | 85800988 (2947, 2948) | Novel Protein sim. GBank gi 2494890 sp Q92176 CORO_BOVIN - CORONIN-LIKE PROTEIN P57 | Contains protein domain (PF00400) - struct WD domain, G-beta repeat | | 264488, 35695917, 35696286, 264692, 33657023, 264693, 33657109, 35696052, 264508, 264905, 264906, 264907, 264629, 264908, 264909, 35696423, 35695855, 264511, 264910, 264632, 264634, 264635, 264636, 264637, 264556, 264557, 264639, 264758, 60432113, 264804, 264605, 264565, 264566, 264764, 264486, 264685, 264766 |
| 1475 | 86871935 (2949, 2950) | | Contains protein domain (PF00041) - Fibronectin type III domain | UNCLASSIFIED | 264681, 264682, 264288, 264566 |
| 1476 | 87548855 (2951, 2952) | Novel Protein sim. GBank gi 4757752 ref NP_004664.1 pANGP - angiotensinogen 3 | Contains protein domain (PF00147) - Fibrinogen beta and gamma chains, C-terminal globular domain | glycoprotein | 60424179, 56181886, 29331824, 60424269, 29331826, 35696052, 264508, 264905, 264906, 264907, 264908, 264909, 264512, 265007, 265008, 265009, 264910, 33657402, 264595, 264596, 55812038, 265011, 264601, 264762, 18108351, 264288, 264369, 264685, 264766, 264689, 55811957, 264691, 264692, 264693, 18108370, 60431528, 18108374, 35696423, 264634, 264635, 264636, 60431850, 264555, 264638, 264557, 264639, 18108382, 18108388, 60432113, 22279002 |
| 1477 | 87774279 (2953, 2954) | Novel Protein sim. GBank gi 2498308 sp Q60870 DP1_MOUSE - POLYPOSIS LOCUS PROTEIN 1 HOMOLOG (TB2 PROTEIN HOMOLOG) (GP106) | | UNCLASSIFIED | 264259, 264107, 264805, 265008, 265010, 265011, 264682, 264288, 265020, 265021, 263974 |
| 1478 | 11754412 (2955, 2956) | | | | 264686 |

| | | | | | |
|------|-----------------------|--|---|--------------|--|
| 1479 | 91640140 (2957, 2958) | Novel Protein sim. GBank gi 5499741 gb AAD43978.1 AF15296 - (AF15296) chromatin-specific transcription elongation factor FACT 140 kDa subunit [Homo sapiens] | | peptidase | 56182575, 22278995, 22278996, 22278998, 22278999, 29331822, 29331824, 66714117, 264906, 264907, 56182435, 265006, 60170831, 33657402, 264758, 33108954, 21906754, 265017, 265019, 264448, 264288, 264767, 264687, 52644229, 21906764, 264689, 21906765, 21906768, 21906769, 265020, 265021, 60170615, 264691, 33657023, 33657109, 33657182, 27486261, 27486262, 33657349, 18108370, 60431528, 263976, 55811576, 264556, 264557, 60170394, 87168518, 264404, 22279000, 22279002, 264563, 264482 |
| 1480 | 94312412 (2959, 2960) | Novel Protein sim. GBank gi 3550456 emb CAA06329.1 (AJ005073) Alix [Mus musculus] | | UNCLASSIFIED | 18108394, 65274572, 56182575, 22278995, 35696286, 56994075, 22278996, 22278997, 22278998, 22278999, 264091, 264259, 35696052, 28148499, 264103, 264105, 264108, 264907, 52644045, 264112, 265007, 265008, 265009, 60433356, 60433438, 264596, 33109954, 33657084, 52644296, 87168474, 265010, 87168559, 265017, 265018, 265019, 264448, 264682, 264683, 264769, 21906765, 21906768, 21906767, 21906768, 21906769, 265020, 265021, 60170615, 52644150, 33657109, 33657182, 263972, 35695855, 264557, 263981, 83373044, 18108385, 87168518, 264566, 264089, 264486 |
| 1481 | 87021442 (2961, 2962) | Novel Protein sim. GBank gi 4836807 gb AAD30566.1 AF14679 - (AF14679) PFT27 [Mus musculus] | | MHC | 265006, 265007, 265010, 18108374 |
| 1482 | 95320442 (2963, 2964) | Novel Protein sim. GBank gi 4585372 gb AAD25403.1 AF12292 - (AF122923) Wnt inhibitory factor-1 [Mus musculus] | Contains protein domain (PF00008) - EGF-like domain | UNCLASSIFIED | 264908, 264910, 264758 |
| 1483 | 94115503 (2965, 2966) | Novel Protein sim. GBank gi 535428 (U13736) - calmodulin- like protein [Pisum sativum] | Contains protein domain (PF00036) - EF hand | struct | 264259, 29331822, 52645080, 29331825, 29331826, 33656970, 29331830, 265007, 55812038, 33109954, 265017, 264288, 21906768, 21906769, 264636, 18108380, 87168518, 22279000 |
| 1484 | 94131544 (2967, 2968) | Novel Protein sim. GBank gi 1911774 bbs 180090 - (S83384) putative Rab5-interacting protein (clone L1-57) [human, HeLa cells, Peptide Partial, 122 aa] [Homo sapiens] | | UNCLASSIFIED | 264489, 35696286, 264259, 264107, 264809, 265008, 60433356, 33657402, 60433438, 264288, 21906765, 21906768, 29148627, 33657023, 27486262, 18108374, 35696423, 83373044, 60432113 |
| 1485 | 80194441 (2969, 2970) | Novel Protein sim. GBank gi 5360129 gb AAD42883.1 AF15511 - (AF155117) NY-REN 62 antigen [Homo sapiens] | Contains protein domain (PF00225) - Kinesin motor domain | struct | 264369, 265020, 18108374 |

| | | | | | |
|------|-----------------------|--|--|---------------|--|
| 1486 | 84125066 (2971, 2972) | Novel Protein sim. GBank gi 4589516 dbj BAA76780.1 - (AB023153) KIAA0938 protein [Homo sapiens] | Contains protein domain (PF00009) - Eukaryotic protein kinase domain | kinase | 56182575, 22278999, 264906, 264907, 21906754, 87168474, 265017, 265019, 18108351, 264288, 265020, 264566, 21906754, 264486 |
| 1487 | 86452711 (2973, 2974) | Novel Protein sim. GBank gi 5019275 emb CAB44431.1 - (AJ132751) xenobiotic/medium-chain fatty acid:CoA ligase form XL-III [Bos taurus] | | synthase | |
| 1488 | 87732026 (2975, 2976) | Novel Protein sim. GBank gi 5712131 gb AAD47378.1 (AF120499) DEM1 protein [Homo sapiens] | Contains protein domain (PF01443) - Viral (Superfamily 1) RNA helicase | tgf | 264686, 264769, 264689, 264692, 264693, 264509, 264906, 264907, 18108370, 264908, 264829, 264909, 264510, 265006, 264512, 265007, 265008, 265009, 264555, 264556, 264557, 264558, 264762, 264564, 264682, 21906767, 22278999, 265022, 264259, 264693, 29331824, 29331825, 29331826, 29331827, 29331828, 264103, 263972, 66712502, 35696423, 35695855, 265007, 265008, 265009, 83373044, 21906754, 56526486, 265017, 264563, 18108351, 264564, 264566, 264369, 264288 |
| 1489 | 85104277 (2977, 2978) | Novel Protein sim. GBank gi 2497303 sp Q82786 FPRP_RAT - PROSTAGLANDIN F2-ALPHA RECEPTOR REGULATORY PROTEIN PRECURSOR (PROSTAGLANDIN F2-ALPHA RECEPTOR ASSOCIATED PROTEIN) | Contains protein domain (PF00047) - Immunoglobulin domain | prostaglandin | 56182575, 264259, 29331822, 29331824, 66714117, 29331827, 29331828, 264508, 264905, 66712502, 265007, 265008, 264594, 33657402, 55812038, 87168474, 265018, 18108351, 264389, 264288, 264769, 264689, 21906767, 21906768, 55811957, 60170615, 33657109, 35695855, 264635, 60170394, 56526486, 22279002, 264563 |
| 1490 | 87390127 (2979, 2980) | | | UNCLASSIFIED | 265007, 264448, 18108372, 264558, 56182323 |
| 1491 | 83594305 (2981, 2982) | Novel Protein sim. GBank gi 295671 (L11275) - selected as a weak suppressor of a mutant of the subunit AC40 of DNA dependant RNA polymerase I and III [Saccharomyces cerevisiae] | | UNCLASSIFIED | |
| 1492 | 85805363 (2983, 2984) | Novel Protein sim. GBank gi 1656005 (U71205) - rt [Mus musculus] | Contains protein domain (PF00071) - Ras family | oncogene | 22278997, 22278998, 29331822, 264907, 66712502 |

| | | | | |
|------|-----------------------|---|------------------|--|
| 1493 | 91677215 (2985, 2986) | Novel Protein sim. GBank gii5689515[dbj BAA83041.1] - (AB029012) KIAA1089 protein [Homo sapiens] | UNCLASSIFIED | 264488, 52646365, 65274572, 56182575, 22278994, 35696286, 56994075, 22278999, 60432049, 29331824, 29331828, 35696052, 264508, 284905, 264908, 52644045, 264909, 56182435, 265008, 265008, 265009, 60170831, 33657402, 55812038, 265010, 265011, 265017, 265018, 265019, 55811150, 264448, 264682, 264685, 264686, 52644229, 21906765, 21906768, 21906768, 21906769, 265020, 265021, 60170615, 52644150, 33657023, 18108384, 18108365, 33657109, 33657182, 27486261, 27486262, 27486264, 33657349, 27486265, 35695763, 18108370, 264629, 18108374, 52644332, 56182323, 87168518, 22279002, 264584, 264566, 264567 |
| 1494 | 87605265 (2987, 2988) | Novel Protein sim. GBank gii728832[sp P39189 ALU2_HUMAN - IIII ALU SUBFAMILY SB WARNING ENTRY IIII] | kinase | 264907, 265009, 264769, 18108370, 55811576, 264639, 264565, 264486 |
| 1495 | 87605267 (2989, 2990) | Novel Protein sim. GBank gii4589588[dbj BAA76816.1] - (AB023189) KIAA0972 protein [Homo sapiens] | transcriptfactor | 22278997, 264259, 264906, 264907, 265009, 264594, 33657084, 265017, 264760, 264448, 33657109, 264630, 264634, 56526486, 264563, 264565, 264568, 264486, 264567 |
| 1496 | 87784322 (2991, 2992) | Novel Protein sim. GBank gii5420387[embj CAB46679.1] - (AJ243459) proteophosphoglycan [Leishmania major] | UNCLASSIFIED | 35696286, 264906, 265019, 264693 |
| 1497 | 81695428 (2993, 2994) | Novel Protein sim. GBank gii3874925[embj CAA92591] - (Z68296) Similarity to Mouse A-RAF proto-oncogene serine/threonine-protein kinase (SW:KRAA_MOUSE); cDNA EST EMBL:D27610 comes from this gene; cDNA EST EMBL:TO1018 comes from this gene; cDNA EST EMBL:D33256 comes from this gene;... | kinase | 264910, 264758, 265011, 264764, 264288, 264690, 264634, 264635, 56526486 |
| 1498 | 80934938 (2995, 2996) | Novel Protein sim. GBank gii728836[sp P39193 ALU6_HUMAN - IIII ALU SUBFAMILY SP WARNING ENTRY IIII] | oncogene | 264488, 65274572, 29331822, 265017, 265018, 21906765, 29148627, 29148629, 18108374, 264637, 264638, 264567 |
| 1499 | 86451589 (2997, 2998) | Novel Protein sim. GBank gii2570198 (U54556) - microfilament sheath protein SHP3 [Litomosoides sigmodontis] | glucoamylase | 263978, 264568 |
| 1500 | 80499386 (2999, 3000) | Novel Protein sim. GBank gii2078483 (U43200) - antifreeze glycopeptide AFGP polypeptide precursor [Boreogadus saida] | UNCLASSIFIED | 22278999, 264769, 18108379 |
| 1501 | 85795297 (3001, 3002) | Novel Protein sim. GBank gii2078483 (U43200) - antifreeze glycopeptide AFGP polypeptide precursor [Boreogadus saida] | UNCLASSIFIED | 264559 |
| 1502 | 80206141 (3003, 3004) | | | 264508, 264112, 264604, 264684, 52644150, 55811576, 264632, 264556, 264638, 56182323, 284563, 264486 |
| 1503 | 87012701 (3005, 3006) | Novel Protein sim. GBank gii3900855 (AC004874) - similar to N-acetylgalactosaminyltransferase; similar to Q07537 (PID.g1171989) [Homo sapiens] | transferase | 29331822, 265007, 264369 |

| | | | | | |
|------|-----------------------|--|--|--------------|--|
| 1504 | 79540051 (3007, 3008) | | Contains protein domain (PF00023) - Ank repeat | UNCLASSIFIED | 264693 |
| 1505 | 86102672 (3009, 3010) | Novel Protein sim. GBank gij475375[emb]CAB41970.1] - (AJ132545) protein kinase [Homo sapiens] | Contains protein domain (PF00069) - Eukaryotic protein kinase domain | kinase | 29331826, 35696052, 264509, 264906, 264907, 264908, 264909, 264511, 264910, 55812038, 264759, 264763, 264764, 264689, 35695917, 265022, 33657109, 18108374, 264631, 264635, 264638, 264566 |
| 1506 | 94143219 (3011, 3012) | Novel Protein sim. GBank gij1304201[dbj]BAA06170] - (D29766) alternatively spliced product [Rattus norvegicus] | Contains protein domain (PF00018) - SH3 domain | glycoprotein | 65274572, 56182575, 56994075, 22278997, 22278998, 22278999, 264081, 264092, 60432049, 264259, 52645080, 29331822, 29331827, 264108, 29331830, 264908, 56182435, 264110, 264511, 264512, 55812038, 21906754, 87168559, 264600, 265017, 265018, 264681, 18108354, 264369, 264687, 264689, 21906765, 29148627, 21906768, 21906769, 29148629, 52644150, 33657023, 18108376, 65274791, 56182323, 264558, 264559, 18108385, 87168518, 60432113, 22279000, 264565, 264639 |
| 1507 | 83738250 (3013, 3014) | Novel Protein sim. GBank gij5689513[dbj]BAA83040.1] - (AB029011) KIAA1088 protein [Homo sapiens] | | helicase | 264593 |
| 1508 | 11618758 (3015, 3016) | Novel Protein sim. GBank gij5031975[ref]NP_005875.1[ppAK4 - protein kinase related to S. cerevisiae STE20, effector for Cdc42Hs] | Contains protein domain (PF00069) - Eukaryotic protein kinase domain | kinase | 264259, 29331822, 29331824, 29331828, 264591, 33109954, 264563 |
| 1509 | 87318451 (3017, 3018) | Novel Protein sim. GBank gij113161[sp]P28614[ACOR - ACETOIN CATABOLISM REGULATORY PROTEIN] | | UNCLASSIFIED | 264259, 29331822, 265007, 18108374, 264556 |
| 1510 | 95362643 (3019, 3020) | Novel Protein sim. GBank gij113161[sp]P28614[ACOR - ACETOIN CATABOLISM REGULATORY PROTEIN] | | UNCLASSIFIED | 264259, 29331822, 29331824, 29331827, 264509, 264907, 264510, 264511, 265007, 264512, 265008, 87168559, 264288, 265022, 33657023, 35695855, 264637, 264638, 264563 |
| 1511 | 88318073 (3021, 3022) | Novel Protein sim. GBank gij128831[sp]P39188[ALU1_HUMAN - III ALU SUBFAMILY J WARNING ENTRY III] | | UNCLASSIFIED | 52645156, 18108396, 56994075, 60432289, 265005, 60433358, 60433438, 21908754, 87168474, 87168559, 265018, 264762, 264763, 264687, 21906765, 21908769, 27486262, 35695763, 18108374, 35696423, 264555, 18108385, 18108387, 18108388, 87168518, 264482 |
| 1512 | 95345390 (3023, 3024) | Novel Protein sim. GBank gij4559353[gb]AAD23014.1[AC00658] putative extragenic suppressor protein [Arabidopsis thaliana] | Contains protein domain (PF01163) - RIO1/ZK632.3/MJ0444 family | | 35696052, 264905, 264906, 264907, 264908, 264909, 264910, 264591, 264766, 264689, 264692, 264629, 264636 |
| 1513 | 87436228 (3025, 3026) | Novel Protein sim. GBank gij1330394 (U58761) - C01F1.6 gene product [Caenorhabditis elegans] | | | |

| | | | | | |
|------|-----------------------|---|--|--------------|---|
| 1514 | 95345392 (3027, 3028) | Novel Protein sim. GBank gi14559353 gb AAD23014.1 AC006585 - (AC006585) putative extragenic suppressor protein [Arabidopsis thaliana] | Contains protein domain (PF01163) - RIO1/ZK632.3/MJ044 family | UNCLASSIFIED | 5264507, 52645156, 52646365, 52646842, 65274572, 22278994, 35698286, 56994075, 284259, 52645080, 29331822, 29331825, 35696052, 29331830, 52644045, 56182435, 265006, 60433358, 60433438, 55812038, 21906754, 52646317, 52644296, 87168474, 87168559, 264448, 52644229, 21906765, 21908766, 21908767, 21908768, 35695917, 265020, 52644150, 33657023, 52645129, 33657109, 33657182, 27486261, 27486262, 27486264, 27486265, 35695763, 18108378, 35696423, 35695855, 52644332, 18108385, 18108387, 87168518, 60432113 265020, 264639 |
| 1515 | 79163538 (3029, 3030) | Novel Protein sim. GBank gi13879501 emb CAA87795 - (Z47812) similar to ubiquitin carboxyl-terminal hydrolase; cDNA EST EMBL:D33368 comes from this gene; cDNA EST EMBL:D33365 comes from this gene; cDNA EST EMBL:D33822 comes from this gene; cDNA EST EMBL:D34547 comes from this ge... | ubiquitin | | |
| 1516 | 88073539 (3031, 3032) | Novel Protein sim. GBank gi1498015 (L27479) - X123 [Homo sapiens] | UNCLASSIFIED | | 265008, 56182323, 22279002 |
| 1517 | 87793325 (3033, 3034) | Novel Protein sim. GBank gi13415134 (AF082024) - Phyb1 [Pimpinella brachycarpa] | | | 264091, 18108370, 264404 |
| 1518 | 87350697 (3035, 3036) | Novel Protein sim. GBank gi1728838 sp P39195 ALU8_HUMAN - IIII ALU SUBFAMILY SX WARNING ENTRY IIII | lm7 | | 66714117, 264508, 264509, 264905, 264510, 264910, 264591, 264595, 264288, 264768, 264789, 18108374, 264636, 264638, 264486 264569, 264489, 60432049, 265009, 33657402, 264598, 21908754, 265019, 264369, 21906765, 21906768, 21908769, 264691, 65274620, 33657182, 27486261, 18108374, 264557, 264639, 87168518, 22279002 |
| 1519 | 94328889 (3037, 3038) | Novel Protein sim. GBank gi15262681 emb CAB45771.1 - (AL080198) hypothetical protein [Homo sapiens] | | | |
| 1520 | 87592855 (3039, 3040) | Novel Protein sim. GBank gi12662161 db BAA23712 - (AB007800) HF0452 cDNA clone for KIAA0440 has a 438- bp insertion at position 1711 of the sequence of KIAA0440. [Homo sapiens] | UNCLASSIFIED | | 18108392, 60432049, 264259, 29331824, 265007, 60433356, 265010, 21906768, 264636 |
| 1521 | 86970696 (3041, 3042) | Novel Protein sim. GBank gi15052351 gb AAD38516.1 AF13542 - (AF13542) GDP- mannose pyrophosphorylase B [Homo sapiens] | Contains protein domain (PF00483) - Nucleotidyl transferase | synthase | 18108394, 264259, 66714117, 265011, 264603, 265019, 18108364, 35696423, 264557, 264558, 18108388 29331824, 265018, 265020, 265021 |
| 1522 | 78960687 (3043, 3044) | Novel Protein sim. GBank gi13776567 (AC005388) - Strong similarity to F2187.33 gi12809264 from A. thaliana BAC gb AC002560. EST gb N65119 comes from this gene. [Arabidopsis thaliana] | UNCLASSIFIED | | |
| 1523 | 91005151 (3045, 3046) | | UNCLASSIFIED | | 55274572, 21906768, 264693 |
| 1524 | 80203723 (3047, 3048) | | UNCLASSIFIED | | 264112, 21906754, 2633974 |
| 1525 | 87799867 (3049, 3050) | Novel Protein sim. GBank gi14759040 ref NP_004283.1 pRIN1 - ras inhibitor | UNCLASSIFIED | | 264683, 264687, 264689, 264690, 264692, 264693 |

| | | | | | |
|------|-----------------------|--|---|--------------|---|
| 1526 | 95105344 (3051, 3052) | Novel Protein sim. GBank gi 7288501sp P08840 AMYH_YEAST - GLUCOAMYLASE S1/S2 PRECURSOR (GLUCAN 1,4-ALPHA- GLUCOSIDASE) (1,4-ALPHA-D-GLUCAN GLUCOHYDROLASE) | | glycoprotein | 35696286, 56182181, 60431735, 264595, 55812038, 264605, 284883, 21906785, 55811957, 265020, 55274791, 264555, 264556, 264557, 264558, 284559, 83373044 |
| 1527 | 88262512 (3053, 3054) | Novel Protein sim. GBank gi 2792496 (AF041107) - tulip 2 [Rattus norvegicus] | | | 56182575, 264259, 60432049, 29331822, 60432289, 264908, 66712502, 60433438, 87168559, 265017, 264288, 21906768, 21906769, 263977, 55811576, 56182323, 18108381 |
| 1528 | 94130918 (3055, 3056) | | | UNCLASSIFIED | 22278995, 22278997, 264259, 66712502, 264596, 265017, 265019, 264682, 264448, 264883, 264764, 264685, 284686, 21906765, 21906766, 21906767, 21906768, 21906769, 265022, 264693, 83373044, 18108385 |
| 1529 | 94120793 (3057, 3058) | Novel Protein sim. GBank gi 406663 gb AAD20053 - (AF131826) Unknown [Homo sapiens] | | UNCLASSIFIED | 264488, 263994, 56182575, 22278995, 35696286, 22278997, 264259, 29331822, 60432289, 29331827, 35696052, 264509, 264906, 264907, 264908, 264909, 52644045, 56182435, 264511, 265009, 264910, 60433356, 60433438, 265017, 265018, 264760, 264448, 264764, 264369, 264288, 264766, 18108357, 264768, 52644229, 21906765, 21906766, 21906767, 21906768, 265021, 265022, 52644150, 33657109, 264629, 35695855, 60432113, 22279002, 264563, 264564, 264486, 264567 |
| 1530 | 95012765 (3059, 3060) | Novel Protein sim. GBank gi 2828710 (AF043642) - maltrin cydophilin [Rattus norvegicus] | | | 264488, 264489, 35696286, 29331825, 35696052, 264508, 264905, 264906, 264907, 264909, 264510, 264511, 264512, 264910, 264592, 264595, 18108351, 264764, 264683, 264684, 264766, 264768, 18108357, 264769, 35695917, 264628, 264629, 18108374, 35695855, 264630, 264631, 264634, 264555, 264636, 264637, 264404, 264563, 264566, 264486 |
| 1531 | 95419351 (3061, 3062) | Novel Protein sim. GBank gi 1905874 (U90878) - carboxyl terminal LIM domain protein [Homo sapiens] | Contains protein domain (PF00595) - PDZ domain (Also known as DHR or GLGF). | kinase | 56182575, 35696286, 264097, 264259, 29331822, 29331825, 29331826, 29331827, 35696052, 264509, 56182435, 264510, 264511, 265007, 60433356, 55811386, 264681, 264369, 264288, 264766, 264687, 55811957, 35695917, 33657023, 35695763, 55810764, 35696423, 55811576, 283981, 60170394, 56182323, 83373044, 60432113, 264568 |

| | | | | | | | |
|------|-----------------------|--|--|--|------------------|--|--|
| 1532 | 85718224 (3063, 3064) | Novel Protein sim. GBank gi 3874716 emb CAA91265 (Z68484) cDNA EST EMBL:D65271 comes from this gene; cDNA EST EMBL:D64845 comes from this gene; cDNA EST EMBL:D64449 comes from this gene; cDNA EST EMBL:D67438 comes from this gene; cDNA EST EMBL:D68087 comes from this gene; cDN... | | | UNCLASSIFIED | 264689 | |
| 1533 | 94239830 (3065, 3066) | Novel Protein sim. GBank gi 1490324 emb CAB01543 (Z78141) unknown [Mus musculus] | | | strud | 29331824, 29146489, 264807, 264112, 265008, 265011, 265017, 265018, 264762, 18108351, 263967, 20281149, 18108374, 263981, 264568 | |
| 1534 | 95343941 (3067, 3068) | Novel Protein sim. GBank gi 81286 pir J22697 - extensin - Vohox carterl (fragment) | | | UNCLASSIFIED | 264905, 264807, 264766, 264637 | |
| 1535 | 90838732 (3069, 3070) | | | | | 65274572, 22278997, 264259, 60432049, 29331822, 60432289, 29331827, 29146499, 265006, 265008, 60170831, 60433438, 33109954, 87168559, 265018, 18108357, 21908768, 29148829, 265021, 265022, 18108377, 56182323, 60432113, 22279000, 22279002 | |
| 1536 | 87602856 (3071, 3072) | Novel Protein sim. GBank gi 106024 pir J332891 - finger protein 2, placental - human | Contains protein domain (PF00096) - Zinc finger, C2H2 type | | transcriptfactor | 264686, 18108357, 18108394, 21906767, 21906768, 29148829, 35696286, 265020, 265021, 52644150, 264693, 68714117, 29331825, 29331826, 264508, 264905, 20281149, 264909, 18108374, 35696423, 35695855, 265009, 264634, 264636, 264638, 18108385, 56526486, 265017, 265018, 264563, 264762, 18108351, 264448, 264369, 264766 | |
| 1537 | 95354556 (3073, 3074) | Novel Protein sim. GBank gi 3876332 emb CAB02086 (Z79754) cDNA EST EMBL:T01054 comes from this gene; cDNA EST EMBL:D73600 comes from this gene; cDNA EST yk426f12.5 comes from this gene; cDNA EST yk342f10.5 comes from this gene; cDNA EST yk475c5.5 comes from this gene; cDNA ES... | | | | 65274572, 56182575, 60432049, 264259, 29331826, 265006, 265007, 60433356, 60433438, 264601, 18108351, 264448, 264389, 264288, 33657023, 65274620, 33657109, 60432113 | |
| 1538 | 85724628 (3075, 3076) | Novel Protein sim. GBank gi 403440 (M81787) - [Gallus domesticus skeletal muscle mRNA, partial cds], gene product [Gallus gallus] | Contains protein domain (PF00069) - Eukaryotic protein kinase domain | | kinase | 18108394, 18108397, 264909, 265008, 265009, 265010, 18108351, 264638, 18108382, 18108385, 18108388 | |

| | | | | | |
|------|-----------------------|--|---|--------------|--|
| 1539 | 95337628 (3077, 3078) | Novel Protein sim. GBank gi 321841 emb CAA19575.1 - (AL023859) SPBC19C7.07c, putative RNA splicing endonuclease ga mima subunit, len:284aa, similar eg. to YAR008W, YAH8_YEAST, P39707, YAR008W, RNA splicing endonuclease gamma subunit, (275aa), (fasta scores, opt:269, E()6.4e-2... | | nuclease | 22278994, 22278996, 35696286, 56994075, 22278997, 22278998, 22278999, 60432049, 264258, 29331822, 29331824, 60424269, 29331825, 60432289, 29331826, 29331827, 29331828, 35696052, 33656970, 56182435, 265009, 33657402, 60433356, 60433438, 55812038, 21906754, 87168559, 265017, 265018, 265019, 18108351, 264288, 52644229, 18108359, 21906764, 21906767, 21906768, 35695917, 265020, 265021, 52644150, 33657023, 33657109, 27486281, 18108370, 18108376, 35696423, 55811576, 65274791, 264558, 56182323, 60170394, 83373044, 87168518, 60432113, 22278000, 22279002, 264586 |
| 1540 | 95352858 (3079, 3080) | Novel Protein sim. GBank gi 5052634 gb AAD38647.1 AF14567 - (AF145672) BCDNA GH12174 [Drosophila melanogaster] | | UNCLASSIFIED | 264369, 264691, 263978 |
| 1541 | 95317948 (3081, 3082) | Novel Protein sim. GBank gi 5052349 gb AAD38515.1 AF13501 - (AF135016) protein phosphatase 2A 48 kDa regulatory subunit [Homo sapiens] | | phosphatase | 264488, 264489, 22278999, 264259, 29331822, 35696052, 264508, 264509, 264905, 264907, 264908, 264511, 264512, 264910, 264592, 264761, 264762, 264448, 264764, 264288, 264687, 21906769, 55811957, 35695917, 265020, 264691, 33657023, 264692, 33657109, 264628, 18108374, 264632, 264634, 264635, 264639, 18108385, 264563, 264564, 264565, 264586, 264486 |
| 1542 | 80937549 (3083, 3084) | Novel Protein sim. GBank gi 5305702 gb AAD41779.1 AF12686 - (AF126867) calpain-like protease [Mus musculus] | | calthepsin | 18108392, 18108394, 65274572, 29331822, 264508, 265007, 265008, 265009, 265011, 264682, 18108354, 18108355, 52644150, 18108368, 264636, 18108381, 18108382 |
| 1543 | 84348768 (3085, 3086) | Novel Protein sim. GBank gi 728832 sp P39189 ALU2_HUMAN - IIII ALU SUBFAMILY SB WARNING ENTRY IJII | | nuclease | 264106, 33109954, 265019, 264683, 35695917, 264690, 264692, 33657109 |
| 1544 | 87757295 (3087, 3088) | Novel Protein sim. GBank gi 3493162 (AF084259) - bromodomain-containing protein BP75 [Mus musculus] | Contains protein domain (PF00439) - Bromodomain | nud_recpt | 35686286, 56994075, 22278999, 35696052, 60433356, 60433438, 265011, 264683, 33657109, 35696423, 264631, 87168518, 22279000 |
| 1545 | 85757973 (3089, 3090) | Novel Protein sim. GBank gi 1086591 (U41007) - similar to S. cerevisiae nuclear protein SNF2 (SP-P22082) in a region of gly-arg repeats [Caenorhabditis elegans] | | UNCLASSIFIED | 264112, 264692, 264693, 55811576 |
| 1546 | 79476589 (3091, 3092) | | | UNCLASSIFIED | 264905, 264686 |
| 1547 | 86999594 (3093, 3094) | Novel Protein sim. GBank gi 2661132 (AF035683) - p21 [Mus musculus] | | UNCLASSIFIED | 264258, 29331822, 66714117, 265007, 55811386, 265010, 264600, 265017, 265019, 264288, 264768, 265020, 265022, 55811578, 18108380, 264563 |

| | | | | | |
|------|-----------------------|--|---|---------------|--|
| 1548 | 94233065 (3095, 3096) | Novel Protein sim. GBank gll3043692[dbj BAA25510] - (AB011156) KIAA0584 protein [Homo sapiens] | | UNCLASSIFIED | 29331824, 60431528, 264639, 56182323 |
| 1549 | 95330048 (3097, 3098) | Novel Protein sim. GBank gll5689519[dbj BAA83043.1] - (AB029014) KIAA1091 protein [Homo sapiens] | | eph | 60424179, 22278995, 35696286, 22278998, 22278999, 264092, 264094, 29331822, 56182181, 29331824, 35696052, 264905, 264906, 264908, 264909, 265006, 264511, 265008, 60431735, 60433356, 21906754, 55811386, 87168559, 265017, 265018, 265019, 55811150, 264682, 264288, 264389, 56181562, 264769, 21906765, 21906769, 21906769, 55811957, 265020, 264691, 33657109, 60431528, 35696423, 35695855, 56526486, 60432113, 22279002, 264563, 264566 |
| 1550 | 95201907 (3099, 3100) | Novel Protein sim. GBank gll544483[sp P35350 GUSB_BOVIN - POSSIBLE GUSTATORY RECEPTOR TYPE B (PPR1 PROTEIN)] | Contains protein domain (PF00001) - 7 transmembrane receptor (rhodopsin family) | Im7 | 65274572, 60432289, 265008, 264910, 265011, 265017, 265019, 264768, 56182323 |
| 1551 | 88077111 (3101, 3102) | Novel Protein sim. GBank gll4758566[ref NP_004798.1 pHS6S - heparan-sulfate 6-sulfotransferase | | UNCLASSIFIED | 22278999, 29331822, 264508, 264509, 264906, 264907, 264909, 265007, 264512, 264910, 21906754, 265018, 265019, 264681, 264764, 264766, 264688, 264769, 21906769, 264692, 35695763, 264635, 264555, 264556, 264557, 264638, 264558, 264563 |
| 1552 | 87617114 (3103, 3104) | | | UNCLASSIFIED | 264259, 29331828, 66712502, 264764, 264288, 264688, 33657109, 264556 |
| 1553 | 94725512 (3105, 3106) | Novel Protein sim. GBank gll4589570[dbj BAA76807.1] - (AB023180) KIAA0963 protein [Homo sapiens] | Contains protein domain (PF00304) - Gamma-thionins family | dehydrogenase | 56182575, 35696286, 29146499, 264509, 264907, 264908, 264909, 56182435, 265006, 265008, 265009, 264910, 264757, 264758, 265017, 55811150, 18108351, 264764, 56181562, 35695917, 264693, 33657109, 18108374, 35696423, 65274791, 35695855, 264635, 264555, 56182323, 18108382, 83373044, 22279000 |
| 1554 | 94233069 (3107, 3108) | Novel Protein sim. GBank gll3043692[dbj BAA25510] - (AB011156) KIAA0584 protein [Homo sapiens] | Contains protein domain (PF00446) - Gonadotropin-releasing hormones | transferase | 35696286, 22278997, 264259, 29331822, 29331824, 29331825, 29331828, 265007, 265009, 60432229, 33657402, 55812038, 265011, 265019, 264681, 264369, 264686, 264767, 264768, 21906765, 21906769, 35695917, 264693, 18108370, 60431528, 55811576, 264631, 60170394, 56182323, 83373044, 18108385, 22279000, 22279002 |

| | | | | | |
|------|-----------------------|---|---|---------------|--|
| 1555 | 87332970 (3109, 3110) | Novel Protein sim. GBank gij2257495[dj]BAA21392] - (AB004534) p1015 [Schizosaccharomyces pombe] | Contains protein domain (PF00400) - WD domain, G-beta repeat | UNCLASSIFIED | 264259, 29331828, 35696052, 264508, 264905, 264906, 264907, 264908, 52644045, 264909, 264910, 60432229, 60433356, 55812038, 264758, 264759, 33657084, 265011, 87168559, 264601, 265018, 265019, 264763, 264764, 264288, 264766, 264769, 21906765, 35695917, 265022, 264691, 33657023, 35696423, 35695855, 264635, 264555, 264636, 264638, 264639, 18108385, 56526486 |
| 1556 | 91228268 (3111, 3112) | | | UNCLASSIFIED | 83373044, 264758, 265022, 264600, 35696052, 264630, 35696423, 265018, 264632, 264682, 29331822, 265020, 265011 |
| 1557 | 87640609 (3113, 3114) | Novel Protein sim. GBank gij3328611 (AF078783) - contains similarity to C3HC4-type zinc fingers (Pfam: zf-C3HC4.hmm, score: 34.08); most similar to drosophila goliath protein (SW: Q06003) [Caenorhabditis elegans] | Contains protein domain (PF00097) - Zinc finger, C3HC4 type (RING finger) | UNCLASSIFIED | 60432289, 264509, 264906, 264907, 264908, 264909, 264910, 264758, 55811386, 264761, 264762, 264766, 264769, 264690, 263978, 264634, 264635, 264639, 264584, 264488 |
| 1558 | 94840376 (3115, 3116) | Novel Protein sim. GBank gij5360105[g]AAD42871.1[AF155105] putative zinc finger protein NY-REN-34 antigen [Homo sapiens] | | | 22278994, 22278996, 22278997, 22278998, 22278999, 60432049, 264259, 29331824, 29331825, 29331826, 29331827, 264908, 264909, 60433356, 21906754, 265017, 265018, 264448, 21906767, 265021, 265022, 33657023, 33657109, 18108370, 55811576, 83373044, 87168518, 22278900, 22279002 |
| 1559 | 88224865 (3117, 3118) | Novel Protein sim. GBank gij112908[sp]P02750[A2GL_HUMAN - LEUCINE-RICH ALPHA-2-GLYCOPROTEIN (LRG)] | Contains protein domain (PF00560) - Leucine Rich Repeat | glycoprotein | 264259, 29331822, 60432289, 35696052, 264107, 264110, 21906754, 33109954, 87168559, 264760, 264763, 21906764, 21906765, 21906769, 265021, 264690, 35695855 |
| 1560 | 94580675 (3119, 3120) | Novel Protein sim. GBank gij3880146[emb]CAA92704] - (Z68319) Similarity to Human hnRNP F protein (PIR Acc. No. S43484); cDNA EST EMBL:D34218 comes from this gene; cDNA EST EMBL:D37248 comes from this gene; cDNA EST EMBL:D71817 comes from this gene; cDNA EST EMBL:D74531 comes from... | | UNCLASSIFIED | 264908, 264603, 264638 |
| 1561 | 86609159 (3121, 3122) | | | UNCLASSIFIED | 264510, 264594 |
| 1562 | 83359682 (3123, 3124) | | | UNCLASSIFIED | 263967 |
| 1563 | 85508694 (3125, 3126) | | | UNCLASSIFIED | 264910, 264764, 264766 |
| 1564 | 87766371 (3127, 3128) | Novel Protein sim. GBank gij1168287[sp]P45953[JACDV_RAT - ACYL-COA DEHYDROGENASE, VERY-LONG-CHAIN SPECIFIC PRECURSOR (VLCD)] | Contains protein domain (PF00441) - Acyl-CoA dehydrogenase | dehydrogenase | 18108394, 35696288, 264259, 29331822, 60432289, 35696052, 29331828, 264508, 66712502, 264908, 56182435, 265007, 264910, 60170831, 21906754, 265011, 265017, 265018, 265019, 18108351, 264448, 264288, 264688, 21906765, 21906768, 265021, 60170615, 264692, 35696423, 35695855, 264557, 56182323, 60432113, 22278902, 264482 |

| | | | | | |
|------|-----------------------|--|--|------------------|--|
| 1565 | 87783381 (3129, 3130) | Novel Protein sim. GBank gii129726[sp]P05307[PDI_BOVIN - PROTEIN DISULFIDE ISOMERASE PRECURSOR (PDI)] (PROLYL 4-HYDROXYLASE BETA SUBUNIT) (CELLULAR THYROID HORMONE BINDING PROTEIN) (P55) | | isomerase | 264488, 264689, 18108398, 55811957, 264534, 264259, 264508, 264905, 264509, 264908, 18108372, 264510, 264511, 264512, 265008, 264630, 265009, 264910, 264635, 264636, 264591, 264555, 264592, 264637, 264593, 264594, 264595, 264596, 265011, 264603, 22279002, 18108351, 264762, 264565, 264567 |
| 1566 | 87424749 (3131, 3132) | Novel Protein sim. GBank gii3880445[emb]CAA20329] - (AL031266) VM106R.1 [Caenorhabditis elegans] | | tnf | 22278996, 22278999, 264259, 29331822, 29331824, 60432288, 29331827, 66712502, 264908, 265008, 18108351, 52644228, 21906765, 21906767, 21906768, 21906769, 33657109, 264555, 264639, 264482 |
| 1567 | 84999006 (3133, 3134) | Novel Protein sim. GBank gii4929699[gb]AAD34110.1[AF15187] CGI-115 protein [Homo sapiens] | | UNCLASSIFIED | 56182575, 21906769, 264692 |
| 1568 | 87648761 (3135, 3136) | Novel Protein sim. GBank gii4827063[ref]NP_005072.1[pZNF1 - zinc finger protein 142 (clone pHZ-49)] | Contains protein domain (PF00096) - Zinc finger, C2H2 type | transcriptfactor | 29331827, 29331830, 264511, 265009, 264758, 21906767, 21908768, 264691, 264693, 22279000, 22279002 |
| 1569 | 90936668 (3137, 3138) | Novel Protein sim. GBank gii5689451[db][BAA83009.1] - (AB028980) KIAA1057 protein [Homo sapiens] | Contains protein domain (PF00443) - Ubiquitin carboxyl-terminal hydrolase family 2 | ubiquitin | 65274572, 29331822, 29331824, 29331828, 264905, 56182435, 265007, 265019, 264764, 21906765, 21906769, 55811957, 60170615, 52644150, 264692, 33657023, 33657109, 18108377, 264563, 264567 |
| 1570 | 86943981 (3139, 3140) | Novel Protein sim. GBank gii1255430 (U53155) - No definition line found [Caenorhabditis elegans] | | UNCLASSIFIED | 264595, 264682, 265021 |
| 1571 | 91210340 (3141, 3142) | Novel Protein sim. GBank gii4507731[ref]NP_001061.1[pTUBG - tubulin, gamma polypeptide] | Contains protein domain (PF00091) - Tubulin/FisZ family | tubulin | 22278996, 35696286, 22278997, 264091, 264259, 29331824, 29331825, 29331827, 35696052, 264508, 264905, 56182435, 264510, 265007, 264758, 265011, 18108351, 264448, 264288, 264369, 21906765, 21906767, 21906768, 21906769, 35695917, 265020, 265021, 33657023, 264693, 18108370, 18108377, 35696423, 35695855, 264634, 264555, 264558, 18108384 |
| 1572 | 81201604 (3143, 3144) | | | UNCLASSIFIED | 35696052, 264905, 264906, 264908, 264910, 264758, 264766, 35695917, 264637 |
| 1573 | 80207068 (3145, 3146) | | | UNCLASSIFIED | 263972 |

| | | | | | |
|------|-----------------------|--|--|--------------|--|
| 1574 | 94216142 (3147, 3148) | Novel Protein sim. GBank gi 4758334 ref NP_004256.1 pFADS - delta-6 fatty acid desaturase | Contains protein domain (PF00173) - Heme-binding domain in cytochrome b5 and oxidoreductases | Cytochrome | 18108394, 264887, 18108397, 18108398, 22278996, 22278997, 22278998, 264259, 29331825, 29331827, 29146498, 29146499, 264107, 264907, 264909, 52644045, 264511, 265008, 264910, 265009, 264591, 21906754, 265011, 265019, 18108351, 264882, 264763, 264764, 18108354, 264369, 264288, 264685, 264766, 264686, 264768, 264688, 21906765, 21906766, 21906767, 21906768, 21906769, 29148628, 264680, 264691, 264693, 20281069, 18108370, 18108374, 18108379, 35695855, 264634, 18108384, 18108385, 22279002, 264563, 264566 |
| 1575 | 95340018 (3149, 3150) | Novel Protein sim. GBank gi 3881810 emb CAA94856 - (Z70783) similar to EF-hand calcium binding protein; cDNA EST EMBL:CO8700 comes from this gene [Caenorhabditis elegans] | Contains protein domain (PF00036) - EF hand | phosphatase | 56994075, 264259, 29331822, 29331824, 29331825, 60432289, 29331828, 264909, 265006, 265008, 265009, 265010, 87168559, 55811150, 264448, 18108354, 264369, 264288, 18108357, 55811957, 265020, 265021, 60170615, 264691, 33657023, 33657109, 60431528, 85274791, 35695855, 18108385, 60432113, 22279002, 264482 |
| 1576 | 95314019 (3151, 3152) | Novel Protein sim. GBank gi 2773195 (AF039711) - contains similarity to Physcomitrella patens glyceraldehyde 3-phosphate dehydrogenase (GB:X72381) [Caenorhabditis elegans] | | UNCLASSIFIED | 264569, 264092, 264094, 264095, 264259, 264508, 264905, 264509, 264907, 264909, 264510, 264511, 265008, 264910, 21906754, 265010, 265011, 87168559, 264761, 264762, 264288, 264766, 264769, 264691, 264693, 35695855, 264632, 264634, 264635, 264638, 83373044, 264486 |
| 1577 | 87613800 (3153, 3154) | Novel Protein sim. GBank gi 2499130 sp P70315 WASP_MOUSE - WISKOTT- ALDRICH SYNDROME PROTEIN HOMOLOG (WASP) | | UNCLASSIFIED | 264511, 265011, 264681, 264369, 264686, 264689, 264629, 264555, 264558, 264559 |
| 1578 | 87123138 (3155, 3156) | | | | 264259, 29331826, 265017, 264689, 264693, 60432113 |
| 1579 | 88085141 (3157, 3158) | Novel Protein sim. GBank gi 2978255 db BAA25190 - (AB007407) myeloid zinc finger protein-2 [Mus musculus] | Contains protein domain (PF00096) - Zinc finger, C2H2 type | UNCLASSIFIED | 35696286, 264908, 264909, 60433438, 55811386, 264369, 264685, 33657023, 264555, 264556, 264557, 87168518 |
| 1580 | 87255702 (3159, 3160) | Novel Protein sim. GBank gi 4324682 gb AAD16986 - (AF109674) late gestation lung protein 1 [Rattus norvegicus] | Contains protein domain (PF00188) - SCP-like extracellular protein | glycoprotein | 22278999, 35696052, 29331830, 52644045, 55812038, 87168474, 265018, 264448, 265022, 264638, 56526486, 22279000 |
| 1581 | 95087431 (3161, 3162) | Novel Protein sim. GBank gi 2088838 (AF003386) - F59E12.4 gene product [Caenorhabditis elegans] | | | 22278995, 29331822, 29331824, 29331826, 56182435, 264595, 55812038, 87168559, 265017, 264288, 21906764, 55811957, 35695917, 264692, 55811576, 264637, 56182323, 264559, 83373044, 60432113 |

| | | | | | |
|------|-----------------------|--|---|---------------|--|
| 1582 | 95358052 (3163, 3164) | Novel Protein sim. GBank gij5420387[embjCAB46679.1] - (A)243459) proteophosphoglycan [Leishmania major] | | phosphatase | 264259, 60432289, 29331827, 264509, 264905, 264906, 264907, 264908, 264910, 264762, 264288, 264768, 264769, 264632, 264555, 264639, 56526486, 22279000 |
| 1583 | 87622715 (3165, 3166) | Novel Protein sim. GBank gij5578959[embjCAB51351.1] - (AL050306) dJ475B7.2 (novel protein) [Homo sapiens] | | UNCLASSIFIED | 60170831, 33657402, 264682, 21906766, 35695855, 264563 |
| 1584 | 95337722 (3167, 3168) | Novel Protein sim. GBank gij5531815[gbjAAD44482.1] - (AF078850) steroid dehydrogenase homolog [Homo sapiens] | Contains protein domain (PF00106) - short chain dehydrogenase | dehydrogenase | 60424179, 52646842, 65274572, 56182575, 22278995, 35696286, 22278996, 22278998, 22278999, 264259, 29331822, 56182181, 60424269, 60432289, 29331827, 29331828, 35696052, 29146498, 66712502, 29331830, 52644045, 56182435, 264510, 264512, 265008, 60433356, 33657402, 60433438, 55812038, 21906754, 55811386, 52644296, 87168474, 87168559, 265018, 265019, 264448, 264369, 264288, 18108358, 21906765, 21906767, 21906768, 21906769, 35695917, 265020, 265021, 265022, 52644150, 33657023, 33657109, 18108374, 55810764, 55811576, 35696423, 65274781, 35695855, 56182323, 83373044, 18108387, 87168518, 60432113, 22278002 |
| 1585 | 87626117 (3169, 3170) | Novel Protein sim. GBank gij4240132[dbjBAA74846.1] - (AB020630) KIAA0823 protein [Homo sapiens] | Contains protein domain (PF00023) - Ank repeat | phosphatase | 35696286, 22278998, 264259, 29331822, 29331824, 29331825, 264905, 265006, 265007, 265008, 60433356, 33109954, 87168474, 265011, 265017, 264604, 264369, 264288, 264685, 264769, 18108359, 21906765, 18108364, 18108370, 264629, 263972, 18108383, 18108388, 264482, 264564 |
| 1586 | 88067081 (3171, 3172) | Novel Protein sim. GBank gij3786494 (AF098993) - No definition line found [Caenorhabditis elegans] | | | 265017, 265018, 264689, 33657023, 263978, 264636, 264563 |
| 1587 | 87617126 (3173, 3174) | Novel Protein sim. GBank gij3253159 (AF005355) - translation initiation factor eIF2C [Oryctolagus cuniculus] | | UNCLASSIFIED | 264907, 264908, 264511, 264910, 264591, 264594, 264629, 264631, 264563, 264483, 264567 |
| 1588 | 87802536 (3175, 3176) | Novel Protein sim. GBank gij1077573[pirjIS2680 - probable ribosomal protein L34, mitochondrial - yeast (Saccharomyces cerevisiae)] | Contains protein domain (PF00468) - Ribosomal protein L34 | UNCLASSIFIED | 264259, 29331828, 264905, 265008, 264758, 21906754, 264761, 284762, 21906765, 21906769, 60170615, 52644150, 33657109, 35695855, 56182323, 18108385 |
| 1589 | 80980653 (3177, 3178) | Novel Protein sim. GBank gij2137756[pirjI48746 - semaphorin C - mouse (fragment)] | | UNCLASSIFIED | 65274572, 264490, 29331822, 66714117, 29331827, 29331828, 56182435, 265008, 60170831, 264595, 264758, 264596, 265011, 264686, 21906766, 21906768, 55811957, 27486265, 264639, 18108385, 56526486, 60432113 |

| | | | | | |
|------|-----------------------|---|---|-------------------|--|
| 1590 | 95319825 (3176, 3180) | | | UNCLASSIFIED | 264489, 22278996, 264259, 29331824, 29331825, 29331826, 29331827, 265006, 60433356, 21906754, 265017, 265018, 265019, 264448, 264785, 264288, 52844229, 21906765, 21906767, 21906768, 21908769, 265021, 264892, 27486285, 35695763, 56526486, 60432113, 22279000, 22279002, 264564 |
| 1591 | 86877160 (3181, 3182) | Novel Protein sim. GBank gi 4557749 ref NP_000237.1 pMHC2 - MHC class II transactivator | | MHC | 264259, 264905, 29331830, 264595, 265017, 264448, 264288, 264690, 264629, 87168518 |
| 1592 | 87682533 (3183, 3184) | | | | |
| 1593 | 94991661 (3185, 3186) | | | UNCLASSIFIED | 65274572, 60432049, 264509, 60433356, 21906754, 21906767, 21906768, 18108370, 35696423, 22279000, 264565, 264567 |
| 1594 | 87773752 (3187, 3188) | Novel Protein sim. GBank gi 3877072 emb CAA87080 - (Z46937) similarity with ribosomal protein L21 [Caenorhabditis elegans] | | UNCLASSIFIED | 264488, 29331827, 264905, 264906, 264907, 264908, 264909, 264910, 264592, 264593, 264757, 264602, 264604, 264760, 264881, 264288, 264766, 264768, 29148629, 35695917, 264692, 264628, 264629, 264630, 264632, 264634, 264635, 264636, 264639, 264563, 264564, 264566 |
| 1595 | 78919425 (3189, 3190) | Novel Protein sim. GBank gi 3152703 (AF065389) - tetraspan NET-4 [Homo sapiens] | Contains protein domain (PF00335) - 4 transmembrane segments integral membrane proteins | UNCLASSIFIED | 29331826, 264908, 55811957 |
| 1596 | 79933928 (3191, 3192) | | | UNCLASSIFIED | 29146498, 264758, 263967 |
| 1597 | 86971857 (3193, 3194) | Novel Protein sim. GBank gi 5257114 gb AAD41244.1 AF09448 - (AF094480) cholesterol 24-hydroxylase [Homo sapiens] | Contains protein domain (PF00067) - Cytochrome P450 | cyto450 | 264092, 29331824, 264508, 264682, 264369, 264686, 264630, 264563 |
| 1598 | 87862939 (3195, 3196) | | | | 264259, 264634 |
| 1599 | 87849829 (3197, 3198) | Novel Protein sim. GBank gi 4506797 ref NP_000324.1 pSCA7 - spinocerebellar ataxia 7 (olivopontocerebellar atrophy with retinal degeneration) | Contains protein domain (PF01581) - FMRFamide related peptide family | UNCLASSIFIED | 52645080, 29331824, 29331826, 264511, 265009, 265011, 264605, 264448, 264764, 265020, 264692, 264693, 18108370, 264635, 18108385 |
| 1600 | 80056002 (3199, 3200) | | | | 29331826, 264603, 264691, 264563 |
| 1601 | 15023246 (3201, 3202) | | | UNCLASSIFIED | 264635 |
| 1602 | 86926987 (3203, 3204) | Novel Protein sim. GBank gi 5305704 gb AAD41780.1 AF12853 - (AF128535) cytoplasmic phosphoprotein PACSIN2 [Mus musculus] | Contains protein domain (PF00018) - SH3 domain | struct | 29148499, 264112, 264762, 18108351, 29148627, 263974 |
| 1603 | 80502072 (3205, 3206) | Novel Protein sim. GBank gi 283920 pr IS27939 - tensin - chicken | | collagen | 264490, 29331824, 264907, 264909, 264511, 265008, 264592, 265010, 265011, 264762, 264764, 264369, 264288, 264687, 264769, 264693, 264628, 264634, 264636, 264555, 264556, 264638, 264557, 264558, 264559, 18108385 |
| 1604 | 80221813 (3207, 3208) | Novel Protein sim. GBank gi 4768831 gb AAD28633.1 AF11682 - (AF116827) unknown [Homo sapiens] | | ATPase-associated | 263977 |

| | | | | | | |
|------|-----------------------|---|---|------------------|-------|--|
| 1605 | 91221129 (3209, 3210) | | | | stuct | 264905, 264509, 264906, 264907, 264908, 264909, 264604, 264768, 264788, 264892, 264683, 33657109, 264629, 35695855, 264635, 264636, 264637 |
| 1606 | 94312703 (3211, 3212) | Novel Protein sim. GBank gi 4505313 ref NP_003794.1 pMYOM - UNKNOWN | Contains protein domain (PF00047) - Immunoglobulin domain | - stuct | | 22278996, 22278999, 264259, 33657402, 265017, 18108351, 264448, 21906767, 21906769, 52644150, 264691, 87168518 |
| 1607 | 10871805 (3213, 3214) | Novel Protein sim. GBank gi 5174473 ref NP_005888.1 pPPI - Intracisternal A particle-promoted polypeptide | | transcriptfactor | | 264689 |
| 1608 | 80428900 (3215, 3216) | Novel Protein sim. GBank gi 2224629 dbj BAA20802 - (AB002342) KIAA0344 [Homo sapiens] | | UNCLASSIFIED | | 264094, 264906, 264907, 264908, 264910, 264591, 264603, 264768, 264693, 264634, 264635, 264637, 264639 |
| 1609 | 94311572 (3217, 3218) | Novel Protein sim. GBank gi 4884073 emb CAB43213.1 - (AL048934) hypothetical protein [Homo sapiens] | | | | 52644507, 52645156, 52646365, 52646842, 56182575, 22278994, 56994075, 35696288, 22278997, 22278998, 22278999, 264259, 52645080, 29147620, 29331828, 35696052, 33656970, 284508, 264509, 264907, 52644045, 56182435, 284510, 264511, 264512, 33657402, 21906754, 52646317, 33109954, 52644296, 87168474, 265017, 265018, 265019, 18108351, 264448, 264288, 264769, 52644229, 21906765, 21906766, 21906767, 21906768, 21906769, 55811957, 35695917, 265020, 265021, 265022, 52644150, 33657023, 33657109, 52645129, 27486261, 27486262, 35695763, 264628, 18108370, 18108376, 35698423, 264638, 52644332, 18108387, 87168518, 22278000, 264563, 264488 |
| 1610 | 85468200 (3219, 3220) | Novel Protein sim. GBank gi 283920 pir J27839 - lensin - chicken | | UNCLASSIFIED | | 264593, 264757, 55812038, 265018, 265020, 264691, 264692, 264693, 264631, 264634, 264635, 264555, 22279000, 264564 |
| 1611 | 94122843 (3221, 3222) | Novel Protein sim. GBank gi 107284 pir J35415 - peroxidase (EC 1.11.1.7), thyroid (Grave's disease) - human (fragment) | Contains protein domain (PF00008) - EGF-like domain | peroxidase | | 35698288, 21906765, 264691, 35696423 |
| 1612 | 85746031 (3223, 3224) | Novel Protein sim. GBank gi 3874846 emb CAA94337 - (Z70307) Similarity to B. subtilis tetracycline resistance protein (SW:TCR2_BACSU); cDNA EST EMBL:CO9951 comes from this gene; cDNA EST EMBL:CO8265 comes from this gene [Caenorhabditis elegans] | | UNCLASSIFIED | | 264488, 264509, 18108370, 18108387, 264486 |
| 1613 | 82247354 (3225, 3226) | | | UNCLASSIFIED | | 264759 |

| | | | | | |
|------|-----------------------|---|--|------------------|--|
| 1614 | 91228634 (3227, 3228) | Novel Protein sim. GBank gi 4880673 gb AAD27726.1 AF13295 - (AF132951) CGI-17 protein [Homo sapiens] | Contains protein domain (PF01605) - eRF1-like proteins | UNCLASSIFIED | 22278995, 22278996, 22278997, 22278998, 22278999, 264259, 29331822, 264908, 264512, 265009, 265011, 265017, 265018, 265019, 18108351, 264683, 264288, 264766, 21906767, 21906768, 21906769, 35695917, 265021, 265022, 35696423, 35695855, 60170394, 56182323, 83373044, 264566 |
| 1615 | 86121909 (3228, 3230) | Novel Protein sim. GBank gi 5689485 dbj BAA83026.1 - (AB028997) KIAA1074 protein [Homo sapiens] | Contains protein domain (PF00023) - Ank repeat | homeobox | 22278996, 35696286, 22278997, 29331822, 35696052, 29331828, 264508, 264908, 264909, 56182435, 264511, 265017, 265019, 264766, 264767, 264768, 265020, 264691, 264628, 264632, 264635, 264555, 264556, 56182323, 264556, 22279002 |
| 1616 | 94311819 (3231, 3232) | Novel Protein sim. GBank gi 3876260 emb CAB01696 - (Z78418) cDNA EST EMBL:D71020 comes from this gene; cDNA EST EMBL:D73593 comes from this gene: cDNA EST EMBL:C07649 comes from this gene: cDNA EST EMBL:C09081 comes from this gene: cDNA EST yk399f2.3 comes from this gene: cDNA ... | | UNCLASSIFIED | 264488, 52644507, 52645156, 52646365, 52646842, 22278994, 22278995, 35696286, 22278996, 22278997, 22278999, 52645080, 29331822, 29331824, 29331825, 29331827, 29331828, 35696052, 33656970, 264905, 264909, 264594, 52646317, 21906754, 33657084, 52644296, 87168474, 87168559, 265017, 265018, 265019, 264681, 264448, 264684, 52644229, 21906764, 264689, 21906765, 21906766, 21906769, 35695917, 265020, 265021, 52644150, 33657023, 52645129, 33657109, 33657182, 27486261, 27486262, 33657349, 27486265, 35695783, 18108376, 35696423, 35695855, 264557, 52644332, 264558, 18108385, 87168518 |
| 1617 | 88090742 (3233, 3234) | Novel Protein sim. GBank gi 466053 sp P34679 Y041_CAEEL - HYPOTHETICAL 68.7 KD PROTEIN ZK757.1 IN CHROMOSOME III | Contains protein domain (PF01529) - DHHC zinc finger domain | peptidase | 35696052, 264905, 264509, 264907, 264908, 264510, 264511, 264764, 264766, 264768, 264889, 264693, 18108374, 264635, 264636, 264638 |
| 1618 | 86272860 (3235, 3236) | Novel Protein sim. GBank gi 4240231 dbj BAA74894.1 - (AB020678) KIAA0871 protein [Homo sapiens] | | struct | 35696286, 22278999, 264092, 29331824, 29331825, 35696052, 33657084, 21906765, 27486264 |
| 1619 | 95354580 (3237, 3238) | Novel Protein sim. GBank gi 5031763 ref NP_005515.1 pHRY - hairy (Drosophila)- homolog | Contains protein domain (PF00010) - Helix-loop-helix DNA-binding domain | transcriptfactor | 52646842, 65274572, 22278999, 264259, 29331822, 29331824, 29331825, 29331826, 29331827, 29331828, 35696052, 56182435, 265007, 265008, 264910, 60170831, 60432229, 60433356, 60433438, 265019, 264448, 264288, 264686, 21906768, 265021, 60170615, 33657023, 65274620, 33657109, 18108374, 18108376, 35696423, 35695855, 56182323, 56526486 |
| 1620 | 87344655 (3239, 3240) | Novel Protein sim. GBank gi 1351047 sp P45843 SCRT_DROME - SCARLET PROTEIN | | UNCLASSIFIED | 264684 |

| | | | | | |
|------|-----------------------|--|---|---------------|--|
| 1621 | 87076708 (3241, 3242) | Novel Protein sim. GBank gi 3822553 (AF098788) - nuclear calmodulin-binding protein [Gallus gallus] | Contains protein domain (PF00622) - SPRY domain | UNCLASSIFIED | 264910 18108392, 65274572, 18108398, 22278996, 22278997, 22278999, 29146498, 29146499, 264905, 264908, 264909, 264828, 52644045, 264592, 60433356, 21906754, 264602, 265017, 264368, 21906768, 55811957, 265021, 60170615, 264635, 264557, 60170394, 83373044, 18108385, 22279000, 22279002, 264566 |
| 1622 | 94741738 (3243, 3244) | Novel Protein sim. GBank gi 3822553 (AF098788) - nuclear calmodulin-binding protein [Gallus gallus] | | UNCLASSIFIED | 18108398, 264259, 264909, 56182435, 87168474, 264448, 21908768, 35695917, 264691, 87168518, 264583 |
| 1623 | 87779106 (3245, 3246) | Novel Protein sim. GBank gi 7310861sp P40389 UV22_SCHPO - UV-INDUCED PROTEIN UV122 | | ribosomalprot | 66714117, 29331825, 264909, 265008, 264758 |
| 1624 | 87338178 (3247, 3248) | Novel Protein sim. GBank gi 3875665 emb CAB05478 - (Z83104) cDNA EST EMBL:100015 comes from this gene; cDNA EST EMBL:D33665 comes from this gene; cDNA EST EMBL:D36540 comes from this gene; cDNA EST yk240f8.3 comes from this gene; cDNA EST yk387c8.3 comes from this gene; cDNA ES... | | UNCLASSIFIED | |
| 1625 | 95354748 (3249, 3250) | Novel Protein sim. GBank gi 4589622 dbj BAA76833.1 - (AB023208) KIAA0989 protein [Homo sapiens] | | kinase | 264489, 22278994, 22278995, 22278996, 35696286, 22278997, 22278998, 22278999, 264092, 264259, 29331824, 29331825, 29331827, 29331828, 264102, 264108, 264508, 33657084, 265017, 265018, 18108351, 264683, 264369, 264288, 21908765, 21908766, 21908767, 21908769, 35695917, 265021, 264691, 65274620, 18108368, 263972, 18108376, 35696423, 264631, 264634, 22279000, 22279002 |
| 1626 | 94734369 (3251, 3252) | Novel Protein sim. GBank gi 5879070 gb AAD46844.1 AF16090 - (AF160904) BcDNA_HL05936 [Drosophila melanogaster] | | | 52644507, 52645156, 52646365, 52646842, 22278994, 56994075, 22278997, 22278998, 22278999, 60432049, 264259, 29331822, 29331824, 66714117, 29331826, 60432289, 29331827, 29331828, 35696052, 264906, 66712502, 264909, 265008, 265009, 60432229, 60433356, 60433438, 21906754, 52646317, 52644296, 265011, 87168559, 264604, 265018, 264448, 264389, 264288, 264766, 52644229, 264689, 21906765, 21906768, 35695917, 265021, 265022, 52644150, 33657023, 65274620, 27486261, 27486262, 27486265, 35695763, 263972, 52644332, 60170394, 87168518, 60432113, 264567 |
| 1627 | 83368773 (3253, 3254) | Novel Protein sim. GBank gi 3668087 (AC004667) - hypothetical protein [Arabidopsis thaliana] | | UNCLASSIFIED | 264288 |
| 1628 | 85708459 (3255, 3256) | Novel Protein sim. GBank gi 3668087 (AC004667) - hypothetical protein [Arabidopsis thaliana] | | eph | 264288, 264686, 264767, 22279002 |

| | | | | | |
|------|-----------------------|---|--|--------------|--|
| 1629 | 84993841 (3257, 3258) | Novel Protein sim. GBank gi 4240175 dbj BAA74866.1 - (AB020850) KIAA0843 protein [Homo sapiens] | | struct | 264555 |
| 1630 | 87779027 (3259, 3260) | | | UNCLASSIFIED | 29331822, 29331827, 265010, 264693, 264634, 22279002 |
| 1631 | 87758454 (3261, 3262) | Novel Protein sim. GBank gi 1915892 emb CAA69995 - (Y08740) tom-1A protein [Gallus gallus] | | UNCLASSIFIED | 55811957, 264259, 33657023, 264693, 29331822, 29331824, 29331827, 29331828, 264906, 264908, 55811576, 264910, 264634, 264636, 264637, 56182323, 264559, 264758, 18108385, 264563, 264764, 264766 |
| 1632 | 87871692 (3263, 3264) | Novel Protein sim. GBank gi 2558501 dbj BAA22896 - (D83850) hepatoma-derived growth factor [Mus musculus] | | UNCLASSIFIED | 264687, 264769, 264691, 264692, 29148499, 264509, 264905, 264907, 264511, 264512, 264482, 264681, 264763, 264682, 264683, 264488, 264259, 264907, 264908, 264909, 264628, 264629, 264631 |
| 1633 | 87773683 (3265, 3266) | | | | 265007, 264637, 22278002 |
| 1634 | 85992817 (3267, 3268) | Novel Protein sim. GBank gi 4887229 gb AAD32244.1 AF15075 - (AF150755) microtubule-actin crosslinking factor [Mus musculus] | Contains protein domain (PF00435) - Spectrin repeat | struct | |
| 1635 | 94232600 (3269, 3270) | | | UNCLASSIFIED | 65274572, 22278998, 35696052, 52644045, 264511, 265008, 265009, 265010, 265011, 265018, 265019, 264448, 264369, 21908765, 21908768, 265021, 264690, 264482 |
| 1636 | 80413227 (3271, 3272) | | | UNCLASSIFIED | 22278995, 264594, 264763, 265020 |
| 1637 | 80070435 (3273, 3274) | Novel Protein sim. GBank gi 455751 ref NP_001339.1 pDAPK - death-associated protein kinase 3 | Contains protein domain (PF00069) - Eukaryotic protein kinase domain | kinase | 264558 |
| 1638 | 87101854 (3275, 3276) | Novel Protein sim. GBank gi 3420051 AC004680 - unknown protein [Arabidopsis thaliana] | | | 21906755, 21906767, 22278998, 35696286, 22278999, 264259, 264692, 264693, 29331824, 33657109, 264508, 264906, 18108370, 264629, 265007, 33657402, 21908754, 264602, 264604, 264764, 264683, 264566, 264288 |
| 1639 | 94322194 (3277, 3278) | Novel Protein sim. GBank gi 5420389 emb CAB46680.1 - (AJ243460) proteophosphoglycan [Leishmania major] | | UNCLASSIFIED | 264488, 18108394, 65274572, 56182575, 35696286, 29331824, 29331826, 29331827, 35696052, 264908, 56182435, 264112, 265006, 265008, 264757, 264758, 55811386, 264603, 264760, 18108351, 264764, 264288, 264766, 264768, 21906767, 55811957, 264691, 33657023, 65274620, 18108370, 55810764, 55811576, 264558, 264639, 83373044, 18108385, 87168518 |

| | | | | | |
|------|-----------------------|--|--|--------------|--|
| 1640 | 94143185 (3279, 3280) | Novel Protein sim. GBank gi 2842469 emb CAA18847.1 - (AL021747) hypothetical protein [Schizosaccharomyces pombe] | | UNCLASSIFIED | 56182575, 56994075, 35696286, 60432049, 60432289, 29331827, 35698052, 52644045, 56182435, 264510, 265006, 265007, 265008, 264910, 265009, 33657402, 55812038, 265010, 265011, 265017, 265018, 264288, 52644229, 21908765, 21908766, 21908768, 35695917, 265021, 60170615, 52644150, 33657023, 33657109, 33657349, 18108374, 35696423, 65274791, 35695855, 264632, 264555, 56182323, 22279000 |
| 1641 | 87625160 (3281, 3282) | | | UNCLASSIFIED | 29148489, 265008, 265007, 265008, 265009, 55812038, 265010, 265011, 264555, 264556, 264558, 18108383 |
| 1642 | 94312557 (3283, 3284) | Novel Protein sim. GBank gi 1575333 (U60416) - myr 6 myosin heavy chain [Rattus norvegicus] | Contains protein domain (PF01843) - struct DIL domain | struct | 22278999, 29147620, 29331826, 29331828, 33659970, 55812038, 265010, 265018, 265019, 18108351, 264689, 265020, 265022, 264690, 33657023, 65274620, 35695763, 52644332, 18108381, 60170394, 58182323, 18108388, 87168518, 22279002, 264564, 29331825, 29331827, 29331828, 21908754, 265019, 264288, 264693, 33657349, 18108370, 18108376, 264555, 83373044, 22279002, 264482 |
| 1643 | 94131766 (3285, 3286) | | | | 264905, 264907, 264908, 264910, 265009, 264757, 264758, 264761, 264762, 264763, 264766, 264768, 264769, 264628, 264629, 264630, 264631, 264632, 264563, 264564, 264565, 264566, 264567 |
| 1644 | 88095125 (3287, 3288) | | | UNCLASSIFIED | 264905, 264907, 264908, 264910, 265009, 264757, 264758, 264761, 264762, 264763, 264766, 264768, 264769, 264628, 264629, 264630, 264631, 264632, 264563, 264564, 264565, 264566, 264567 |
| 1645 | 95013858 (3289, 3290) | | | UNCLASSIFIED | 264685, 264693 |
| 1646 | 95362691 (3291, 3292) | Novel Protein sim. GBank gi 1076802 pir IS49915 - extensin like protein - maize | | UNCLASSIFIED | 22278994, 56994075, 35696286, 264259, 29331824, 29331825, 29331826, 60432289, 264508, 60433356, 60433438, 87168559, 265018, 264687, 35695917, 264692, 33657023, 33657182, 27486261, 27486265, 33657349, 60432113, 264563, 264564 |
| 1647 | 94278428 (3293, 3294) | Novel Protein sim. GBank gi 5002573 emb CAB44338.1 - (Y17466) alpha-N-acetylgalactosamine alpha-2,6-sialyltransferase [Fugu rubripes] | | UNCLASSIFIED | 29331822, 264906, 264908, 264369, 21908768, 60170615, 264639, 22279000 |
| 1648 | 87642098 (3295, 3296) | | Contains protein domain (PF00096) - Zinc finger, C2H2 type | UNCLASSIFIED | 265009, 264686, 55811957, 35695917, 55810764, 264556, 56182323, 264558, 18108385 |

| | | | | | |
|------|-----------------------|---|--|--------------|--|
| 1649 | 95347628 (3297, 3298) | Novel Protein sim. GBank gij1854065[emb CAA58337] - (X83413) U88 [Human herpesvirus 6] | | cadherin | 264488, 22278995, 35696286, 22278998, 22278997, 22278999, 60432049, 264259, 29331822, 29331824, 29331825, 29331826, 29331827, 29331828, 264905, 264907, 66712502, 264908, 52644045, 264909, 56182435, 264511, 265007, 265008, 265009, 264591, 264593, 60433438, 264596, 55812038, 21906754, 265011, 264601, 264602, 265017, 265018, 265019, 264682, 264448, 264764, 264683, 264288, 264768, 264685, 264687, 264768, 264688, 264769, 52644229, 264689, 21906765, 21906766, 21906767, 21906768, 55811957, 35695917, 265021, 265022, 52644150, 264692, 33657023, 33657109, 20281149, 18108370, 264628, 18108374, 18108376, 35696423, 35695855, 264632, 264634, 264635, 264636, 18108380, 264639, 264556, 18108382, 18108384, 18108385, 18108387, 264080, 264404, 60432113, 22279000, 22279002, 264482, 264555, 264566, 264487, 265011, 264602, 21906767, 18108374, 18108377, 18108385 |
| 1650 | 87418539 (3299, 3300) | Novel Protein sim. GBank gij3647335[emb CAA21059] - (AL031644) possible zinc-finger protein [Schizosaccharomyces pombe] | | | |
| 1651 | 91639773 (3301, 3302) | Novel Protein sim. GBank gij4884278[emb CAB43247.1] - (AL050037) hypothetical protein [Homo sapiens] | | synthase | 264488, 52645156, 18108397, 35696286, 22278998, 22278999, 264259, 29331822, 29331824, 29331825, 29331826, 29331827, 29331828, 264508, 264908, 29331830, 264910, 60432229, 21906754, 265010, 265011, 265017, 265019, 264448, 18108354, 264288, 264688, 21906765, 21906766, 21906768, 21906769, 265022, 264692, 264693, 264629, 35695855, 264556, 264637, 264557, 264559, 83373044, 56526486, 22279000, 22279002, 264564 |
| 1652 | 86598622 (3303, 3304) | Novel Protein sim. GBank gij1657837 (U73200) - p116Rip [Mus musculus] | Contains protein domain (PF00169) - PH domain | struct | 22278997, 29146498, 56182435, 21906754, 264369, 21906765, 21906768, 21906769, 265020, 52644150, 33657109, 22279000, 22279002 |
| 1653 | 94255993 (3305, 3306) | Novel Protein sim. GBank gij3776054[emb CAA06273] - (AJ004999) Tapasin [Gallus gallus] | Contains protein domain (PF00047) - glycoprotein Immunoglobulin domain | | 18108398, 22278995, 22278998, 264259, 29331822, 29331824, 60432289, 29331826, 29331827, 29331830, 264909, 265006, 265009, 60432229, 60433356, 60433438, 21906754, 265017, 265019, 264448, 264683, 264288, 265021, 265022, 264692, 18108364, 55274791, 18108384, 60432113, 264567, 33657109, 264565 |
| 1654 | 79756471 (3307, 3308) | | | UNCLASSIFIED | |

| | | | | | |
|------|-----------------------|---|--|--------------|--|
| 1655 | 86689346 (3309, 3310) | Novel Protein sim. GBank gij3355717[embjCAA73496] - (Y13053) seryl-tRNA synthetase [Zea mays] | | synthase | 52644507, 35696286, 22278998, 22278999, 29331824, 29331825, 29331828, 33656970, 264908, 52644045, 264511, 264910, 52646317, 264288, 52644229, 33657023, 33657109, 52644332, 264557, 56182323, 56526486, 80432113 |
| 1656 | 79962297 (3311, 3312) | Novel Protein sim. GBank gij1890141[dbjBAA18947] - (D83206) P24 protein [Mus musculus] | | UNCLASSIFIED | 29331822, 29331824, 29331825, 264563 |
| 1657 | 87771994 (3313, 3314) | Novel Protein sim. GBank gij4557645[refjNP_001524.1]pHNRP - heterogeneous nuclear ribonucleoprotein L | Contains protein domain (PF00076) - RNA recognition motif (a.k.a. RRM, RBD, or RNP domain) | dna_ma_bind | 29331827, 265009, 21906766, 21906767, 265020, 265022, 33657109, 264638, 56526488, 264482 |
| 1658 | 87773778 (3315, 3316) | Novel Protein sim. GBank gij3877072[embjCAA87060] - (Z48937) similarity with ribosomal protein L21 [Caenorhabditis elegans] | Contains protein domain (PF00829) - Ribosomal prokaryotic L21 protein | UNCLASSIFIED | 52646365, 35696286, 22278996, 22278997, 22278998, 22278999, 264259, 29331822, 29331824, 29331825, 29331826, 29331827, 29331828, 29146498, 264905, 264908, 52644045, 265006, 60433356, 264757, 60433438, 21906754, 265011, 18108351, 264448, 264369, 264288, 264766, 264768, 21906765, 21906767, 21906768, 21906769, 29148629, 265021, 265022, 18108362, 263969, 263971, 18108374, 35696423, 18108383, 22279000, 264482 |
| 1659 | 88230101 (3317, 3318) | Novel Protein sim. GBank gij539218[pirjS38038] - hypothetical protein YKL201c - yeast (Saccharomyces cerevisiae) | | UNCLASSIFIED | 52646317, 21906766, 21906767, 21906768, 87168518, 22278996, 265020, 22278999, 87168559, 264603, 265017, 264631, 265018, 265019, 22279002, 264482, 264635, 264565, 264488, 35696286, 264259, 35696052 |
| 1660 | 94315313 (3319, 3320) | Novel Protein sim. GBank gij2497012[spjQ10010]YSV4_CAEEL - HYPOTHETICAL 26.6 KD PROTEIN T19C3.4 IN CHROMOSOME III | | UNCLASSIFIED | 264508, 264509, 264905, 264906, 264907, 264908, 264909, 264510, 264511, 265006, 265007, 264512, 265009, 264910, 264592, 264598, 265010, 264600, 264602, 265017, 265018, 264605, 264760, 264764, 264288, 264766, 264688, 264768, 264769, 264689, 21906766, 35695917, 264690, 33657023, 264693, 33657109, 264629, 35696423, 35695855, 264634, 264635, 264555, 264636, 264637, 264556, 264638, 264639, 264559, 18108385, 18108388, 264563, 264483, 264564, 264565, 264566, 264486, 264567 |

| | | | | | |
|------|-----------------------|--|---|--------------|---|
| 1661 | 94234071 (3321, 3322) | Novel Protein sim. GBank gi 4759100 ref NP_004759.1 pSFRS - splicing factor, arginine/serine-rich 11 | Contains protein domain (PF00076) - RNA recognition motif. (a.k.a. RRM, RBD, or RNP domain) | UNCLASSIFIED | 264488, 22278998, 264259, 29331824, 29331826, 29331827, 29331828, 264509, 66712502, 29331830, 264908, 52644045, 265007, 264512, 60433356, 60433438, 55812038, 21906754, 265019, 264448, 264766, 264768, 264769, 21908768, 21906769, 265020, 33657023, 33657109, 65274791, 87168518, 264482, 264563, 264564, 264565, 264567 |
| 1662 | 94135172 (3323, 3324) | Novel Protein sim. GBank gi 1730502 sp P52875 PF27_MOUSE - TRANSMEMBRANE PROTEIN PF27 | | | 18108392, 29331822, 29331828, 20281100, 264106, 265006, 265007, 265008, 18108348, 21906766, 18108365, 18108366, 18108374, 83373044, 18108385 |
| 1663 | 94217146 (3325, 3326) | Novel Protein sim. GBank gi 4884136 emb CAB43275.1 - (AL050107) hypothetical protein [Homo sapiens] | Contains protein domain (PF00397) - WW domain | kinase | 52645156, 56182575, 22278994, 22278995, 35696286, 22278996, 56994075, 22278997, 22278998, 22278999, 264259, 29331822, 29331826, 29331827, 29331828, 33655970, 29331830, 264908, 56182435, 264511, 60433356, 33657402, 33109954, 87168474, 87168559, 265017, 265018, 264605, 18108351, 264764, 264288, 264766, 264768, 21906765, 21908766, 21908767, 21908768, 21908769, 265021, 265022, 264691, 33657023, 264693, 263987, 33657109, 264630, 52644332, 83373044, 87168518, 80432113, 22279000 |
| 1664 | 94234076 (3327, 3328) | Novel Protein sim. GBank gi 3043682 dbj BAA25510 - (AB011156) KIAA0584 protein [Homo sapiens] | | UNCLASSIFIED | 264488, 263994, 35696286, 29331824, 35696052, 264508, 264509, 264905, 264906, 264907, 264908, 264909, 264510, 264511, 265009, 264910, 60170831, 264591, 264592, 264595, 87168474, 265011, 264600, 264601, 264604, 264605, 264760, 264762, 18108351, 264681, 264682, 264763, 264683, 264764, 264288, 264684, 264766, 264687, 264768, 264769, 21906764, 21908765, 21908767, 35695917, 265021, 264534, 60170615, 264690, 264691, 264692, 33657109, 33657182, 264628, 18108370, 264629, 35698423, 35695855, 264634, 264635, 264555, 264636, 264637, 264638, 264639, 264558, 83373044, 87168518, 264563, 264566, 264486 |

| | | | | | |
|------|-----------------------|---|---|--------------|--|
| 1665 | 91226852 (3329, 3330) | Novel Protein sim. GBank gjl1083506[ipr]S50065 - sialoadhesin - mouse | Contains protein domain (PF00047) - Immunoglobulin domain | immunoglob | 264488, 29331826, 29331828, 264509, 264906, 264907, 264909, 264510, 264511, 264910, 264592, 264593, 264595, 264758, 264596, 264600, 264760, 264762, 264764, 264766, 264768, 264629, 264630, 264634, 264636, 83373044, 264564, 264566, 264567, 264486 |
| 1666 | 95358160 (3331, 3332) | Novel Protein sim. GBank gjl3913431[sp]O42643DDX8_SCHPO - PUTATIVE PRE-MIRNA SPLICING FACTOR ATP-DEPENDENT RNA HELICASE SPAC10F6.02C | Contains protein domain (PF00575) - S1 RNA binding domain | helicase | 58994075, 22278999, 264259, 29331824, 29331826, 29331827, 29146498, 265009, 33109954, 87168559, 265019, 264288, 264688, 21906767, 21906769, 264691, 33657182, 18108370, 18108374, 18108385, 22279002 |
| 1667 | 91226655 (3333, 3334) | Novel Protein sim. GBank gjl5689535[dbj]BAAB3051.1] - (AB029022) KIAA1099 protein [Homo sapiens] | Contains protein domain (PF01412) - Putative GTP-ase activating protein for Arf | transport | 264259, 29331822, 29331826, 264905, 264906, 264908, 264510, 265009, 264595, 264758, 265011, 87168559, 265017, 265018, 265019, 264448, 264768, 264686, 21906765, 21906767, 21906769, 265020, 265021, 60170615, 264690, 264692, 264693, 18108368, 18108370, 263972, 55810764, 264555, 83373044, 60432113, 22279000, 22279002 |
| 1668 | 88095135 (3335, 3336) | Novel Protein sim. GBank gjl2076894[gb]AAB53983.1] - (AF002197) short region of weak similarity to protein kinase C; contains similarity to Pfam domain PF00130 (DAG_PE_bind). Score=10.0, E-value=0.0034, N=1 [Caenorhabditis elegans] | | kinase | 66714117, 264508, 264509, 264906, 264907, 264908, 264511, 264910, 264764, 264687, 264689, 33657109, 35695423, 35695855, 264632 |
| 1669 | 91227848 (3337, 3338) | Novel Protein sim. GBank gjl387537[emb]CAA85414.1] - (Z36948) contains a valine and arginine rich domain, possesses weak similarity with the RNA binding domains from RNA splicing factor UZAF 65 KD subunit; cDNA EST EMBL:D64658 comes from this gene; cDNA EST EMBL:D66829 comes fr... | | UNCLASSIFIED | 29331825, 33109954, 264369, 264767, 264689, 33657109, 83373044 |
| 1670 | 87628009 (3339, 3340) | | | UNCLASSIFIED | 264259, 29331824, 29331827, 60433438, 265022, 264636 |
| 1671 | 87346372 (3341, 3342) | Novel Protein sim. GBank gjl462451[sp]P34244[KKK1_YEAST - PROBABLE SERINE/THREONINE-PROTEIN KINASE YKL101W | Contains protein domain (PF00069) - Eukaryotic protein kinase domain | kinase | 35696286, 22278997, 29331825, 264909, 21906754, 265017, 265018, 265019, 264682, 264683, 264768, 264688, 21908768, 21906767, 21906768, 21906769, 264691, 264555, 264556, 22279000, 264566 |
| 1672 | 88291834 (3343, 3344) | Novel Protein sim. GBank gjl1814270 [U74586] - double-stranded RNA specific adenosine deaminase [Rattus norvegicus] | | deaminase | 264906, 264909, 264632, 18108381 |

| | | | | | |
|------|-----------------------|---|--|--------------|--|
| 1673 | 88095137 (3345, 3346) | Novel Protein sim. GBank gi 2076894 gb AAB53983.1 - (AF002197) short region of weak similarity to protein kinase C; contains similarity to Pfam domain PF00130 (DAG_PE-bind). Score=10.0, E-value=0.0034, N=1 [Caenorhabditis elegans] | Contains protein domain (PF00130) - Phorbol esters/diacylglycerol binding domain (C1 domain) | kinase | 264488, 264568, 18108394, 56994075, 22278996, 264259, 35696052, 264508, 264905, 264509, 264906, 264907, 264908, 264909, 264510, 264910, 60170831, 264592, 264594, 264595, 264758, 264601, 264780, 264762, 264683, 264764, 264288, 264766, 264686, 264768, 264687, 264769, 264689, 264680, 33657023, 264692, 264693, 33657109, 264628, 264629, 18108374, 35696423, 35695855, 264631, 264632, 264634, 264635, 264637, 264556, 264638, 264639, 264563, 264482, 264564, 264565, 264566, 264567, 264486 |
| 1674 | 88258028 (3347, 3348) | Novel Protein sim. GBank gi 5262467 emb CAB45693.1 - (AL080062) hypothetical protein [Homo sapiens] | | kinase | 29331822, 29331824, 264906, 52644045, 60433358, 87168559, 264448, 264288, 264888, 264691 |
| 1675 | 87606466 (3349, 3350) | Novel Protein sim. GBank gi 3128366 (AF010496) - 50S ribosomal protein l9 [Rhodobacter capsulatus] | | UNCLASSIFIED | 56181686, 35696286, 22278997, 22278998, 264259, 29331824, 29331827, 35696052, 66712502, 264764, 264288, 264686, 264687, 35695917, 265020, 264690, 264693, 35695763, 18108370, 35696423, 35695855, 264637, 264639, 18108385, 264584 |
| 1676 | 95358086 (3351, 3352) | Novel Protein sim. GBank gi 4164065 gb AAD05327 - (AF111091) latrophilin 3 splice variant bba1 [Bos taurus] | | UNCLASSIFIED | 264259, 29331827, 29331828, 264106, 264907, 265009, 264600, 265019, 264288, 21906765, 265020, 265022, 35695855, 83373044, 18108385 |
| 1677 | 87408587 (3353, 3354) | Novel Protein sim. GBank gi 3327046 dbj BAA31591 - (AB014516) KIAA0616 protein [Homo sapiens] | | UNCLASSIFIED | 264908 |
| 1678 | 86866829 (3355, 3356) | Novel Protein sim. GBank gi 550452 (U08469) - 3-methylcrotonyl-CoA carboxylase, biotin-carrier domain [Glycine max] | Contains protein domain (PF00289) - Carbamoyl-phosphate synthase (CPSase) | UNCLASSIFIED | 29331824, 264102 |
| 1679 | 91214108 (3357, 3358) | Novel Protein sim. GBank gi 550452 (U08469) - 3-methylcrotonyl-CoA carboxylase, biotin-carrier domain [Glycine max] | | carboxylase | 264488, 18108392, 18108394, 52646842, 18108397, 18108398, 35696286, 29331824, 265006, 265007, 265008, 265009, 18108348, 265011, 18108351, 264883, 18108354, 18108358, 18108359, 21906765, 29148627, 29148629, 264690, 18108361, 18108362, 18108364, 18108365, 18108368, 264628, 18108379, 35696423, 35695855, 264635, 18108381, 18108382, 18108383, 18108384, 18108385, 18108388 |
| 1680 | 81005372 (3359, 3360) | Novel Protein sim. GBank gi 2394478 (AF024500) - No definition line found [Caenorhabditis elegans] | | transport | 65274572, 22278994, 22278999, 66714117, 29331827, 56182435, 21806754, 265018, 264288, 21906769 |
| 1681 | 94324150 (3361, 3362) | Novel Protein sim. GBank gi 5689537 dbj BAA83052.1 - (AB029023) KIAA1100 protein [Homo sapiens] | | UNCLASSIFIED | 22278996, 29331822, 264908, 264593, 264604, 265019, 264683, 55811957, 264690, 33657023, 35696423, 83373044, 264563 |
| 1682 | 86042710 (3363, 3364) | | | UNCLASSIFIED | 264909, 265017, 264605 |

| | | | | | |
|------|-----------------------|---|---|--------------|--|
| 1683 | 94316213 (3365, 3366) | Novel Protein sim. GBank gi 5031717 ref NP_005704.1 pGPBP - goodpasture antigen- binding protein | Contains protein domain (PF01852) - START domain | UNCLASSIFIED | 263994, 35696286, 35696052, 264508, 264509, 264905, 264906, 264907, 264908, 264909, 264511, 265006, 265007, 265009, 264910, 264593, 264756, 265010, 265011, 264760, 264761, 264762, 264763, 264764, 264369, 264766, 264768, 35695917, 264692, 33657109, 35695423, 35695855, 264634, 264635, 264636, 264638, 264639, 83373044, 264486 |
| 1684 | 80063409 (3367, 3368) | | | UNCLASSIFIED | 264563, 264566 |
| 1685 | 94323182 (3368, 3370) | Novel Protein sim. GBank gi 1255371 (U53147) - coded for by C. elegans cDNA yk34a9.5; coded for by C. elegans cDNA yk34a9.3; Similar to guanylate kinase. [Caenorhabditis elegans] | Contains protein domain (PF00625) - Guanylate kinase | UNCLASSIFIED | 60424179, 52646842, 22278994, 35696286, 22278998, 264259, 52645080, 29331824, 29331826, 265007, 33657084, 265018, 264681, 264448, 264683, 264369, 264689, 21906765, 21906767, 21906768, 21906769, 265021, 264692, 65274620, 33657109, 27486262, 264635, 52644332, 56182323, 22279000 |
| 1686 | 87820710 (3371, 3372) | Novel Protein sim. GBank gi 2244707 dbj BAA21115.1 - (AB005287) thrombospondin 1 [Bos taurus] | | UNCLASSIFIED | 56182575, 264259, 264508, 264905, 264509, 264907, 264908, 264510, 264511, 265006, 264512, 265009, 264910, 264758, 265010, 265011, 264605, 18108351, 264764, 264766, 18108357, 264768, 18108362, 264628, 264630, 264631, 264634, 264635, 264637, 264638, 264639, 264655, 264486, 264567 |
| 1687 | 94719400 (3373, 3374) | Novel Protein sim. GBank gi 4680679 gb AAD27729.1 AF13295 - (AF132954) CGI-20 protein [Homo sapiens] | | UNCLASSIFIED | 35659517, 264906, 264907, 264908, 264510, 265006, 265007, 264910, 264556, 18108381, 18108383, 265011 |
| 1688 | 82158442 (3375, 3376) | | | UNCLASSIFIED | 264905, 264910, 264760, 264629, 264555 |
| 1689 | 94325049 (3377, 3378) | Novel Protein sim. GBank gi 4240193 dbj BAA74875.1 - (AB020659) KIAA0852 protein [Homo sapiens] | | UNCLASSIFIED | 264559, 35696286, 22278999, 264508, 264908, 264909, 56182435, 33657402, 60433438, 55812038, 265017, 265018, 264448, 264764, 264288, 264686, 29148629, 35695917, 265020, 265021, 263972, 18108374, 65274791, 83373044, 264089 |
| 1690 | 83255346 (3379, 3380) | Novel Protein sim. GBank gi 3800736 (AF031572) - seven- pass transmembrane receptor precursor [Mus musculus] | Contains protein domain (PF00028) - Cadherin domain | cadherin | 264369, 21906766, 264692, 264639, 87188518 |
| 1691 | 88095223 (3381, 3382) | Novel Protein sim. GBank gi 2773208 (AF039713) - No definition line found [Caenorhabditis elegans] | | | 264768, 33657109, 29331827, 29148629, 264510, 264106, 264910, 264109, 264508, 60170831, 264563, 264905, 264564, 264691, 264637, 264628, 264807, 264908, 33657023, 264567, 264766, 263974 |
| 1692 | 86106709 (3383, 3384) | | | | 264106 |
| 1693 | 87012775 (3385, 3386) | Novel Protein sim. GBank gi 121271 sp P02207 GLB_LAMFL - GLOBIN | Contains protein domain (PF00042) - Globin | UNCLASSIFIED | 29331826, 264508, 264905, 264907, 264595, 265010, 265011, 21906768, 33657023, 264629, 263978, 264558 |

| | | | | | |
|------|-----------------------|--|--|--------------|--|
| 1694 | 94208168 (3387, 3388) | Novel Protein sim. GBank gi 5453932 ref NP_006225.1 pPOLR - polymerase (RNA) II (DNA directed) polypeptide J (13.3kD) | Contains protein domain (PF01193) - RNA polymerases L / 13 to 16 kDa subunit | mapolymerase | 35698286, 22278998, 22278998, 22278998, 264259, 29331822, 29331825, 29331826, 29331828, 35696052, 29146499, 264905, 264908, 52844045, 264511, 265006, 265007, 265009, 264592, 60433356, 21906754, 265010, 265011, 18108351, 264763, 264682, 264448, 264683, 264288, 264768, 264689, 21906768, 60170615, 264691, 264692, 264693, 18108370, 18108374, 263978, 35696423, 35695855, 264556, 18108381, 18108385, 87168518, 264482, 264486 264634 |
| 1695 | 94719325 (3389, 3390) | Novel Protein sim. GBank gi 4680679 gb AAD27729.1 AF13295 - (AF132954) CGI-20 protein [Homo sapiens] | | UNCLASSIFIED | |
| 1696 | 87824038 (3391, 3392) | Novel Protein sim. GBank gi 4220517 emb CAA22990 - (AL035356) hypothetical protein [Arabidopsis thaliana] | | | 22278997, 264259, 265010, 18108351, 264764, 21906766, 18108370 |
| 1697 | 85740963 (3393, 3394) | Novel Protein sim. GBank gi 505662 (U10362) - GP36b glycoprotein [Homo sapiens] | | glycoprotein | 264682 |
| 1698 | 87445285 (3395, 3396) | Novel Protein sim. GBank gi 5052031 gb AAD38411.1 AF15573 - (AF155739) axolotrin [Mus musculus] | | | 56994075, 22278998, 22278999, 264509, 33657402, 264758, 87168474, 87168559, 265017, 265018, 264448, 264687, 29148627, 21906769, 29148629, 265020, 265022, 33657023, 264558, 87168518, 22279002 |
| 1699 | 87424793 (3397, 3398) | | | UNCLASSIFIED | 35696286, 264635 |
| 1700 | 87859161 (3399, 3400) | Novel Protein sim. GBank gi 543344 pir S41647 - zinc finger 5 protein - mouse | Contains protein domain (PF00096) - Zinc finger, C2H2 type | nucl_rept | 29331824, 52644045, 265008, 265009, 263969, 263971 |
| 1701 | 86570488 (3401, 3402) | | | UNCLASSIFIED | 263977 |
| 1702 | 87795092 (3403, 3404) | Novel Protein sim. GBank gi 3877439 emb CAA96652 - (Z72510) similarity to yeast UTR3 protein (Swiss Prot accession number P21374); cDNA EST EMBL:D72822 comes from this gene; cDNA EST EMBL:D75763 comes from this gene; cDNA EST yk274e3.3 comes from this gene; cDNA EST yk274e3.5 c... | | MHC | 22278995, 22278997, 264092, 29146498, 29146499, 264107, 264508, 264907, 264110, 264112, 265009, 60170831, 21906754, 265011, 265017, 264762, 18108351, 264288, 21906765, 35695917, 265021, 60170815, 263967, 33657109, 18108370, 263972, 263974, 18108374, 263976, 35695855, 264555, 263981, 60170394, 18108385, 56528486, 87168518, 60432113 |
| 1703 | 79588651 (3405, 3406) | Novel Protein sim. GBank gi 451544 (U04267) - proline-rich cell wall protein [Gossypium barbadense] | | UNCLASSIFIED | 264909, 265017, 264628, 264629, 264638 |
| 1704 | 86622978 (3407, 3408) | Novel Protein sim. GBank gi 1263289 (U47856) - fibroin-4 [Araneus diadematus] | | UNCLASSIFIED | 264369 |
| 1705 | 87795175 (3409, 3410) | Novel Protein sim. GBank gi 4519621 db BAA75670.1 - (AB017614) OASIS protein [Mus musculus] | | | 264569, 35696286, 264907, 265010, 264687, 264768, 264692, 264693, 264636, 264566 |
| 1706 | 87790967 (3411, 3412) | Novel Protein sim. GBank gi 3123034 sp Q15011 Y025_HUMAN - HYPOTHETICAL PROTEIN KIAA0025 | | | 22278996, 22278998, 264259, 264509, 265018, 264764, 264685, 264686, 21906768, 21906769, 265022, 264691, 264558, 22279000 |

| | | | | | |
|------|-----------------------|--|--|------------------|--|
| 1707 | 88041230 (3413, 3414) | Novel Protein sim. GBank gll4321684[gb AAD15797] - (AF055470) ZNF258 [Homo sapiens] | | UNCLASSIFIED | 18108396, 22278997, 264259, 29147620, 29331826, 29146498, 264905, 264906, 265008, 264593, 264595, 264758, 264596, 265018, 264760, 18108351, 264764, 264766, 264689, 264693, 18108370, 35698423, 55811576, 264558, 87168518, 60432113, 264567 |
| 1708 | 91220519 (3415, 3416) | Novel Protein sim. GBank gll517459[ref NP_005947.1 pMTHF - 5,10-methylenetetrahydrofolate dehydrogenase, 5,10-methylenetetrahydrofolate cyclohydrolase, 10-formyltetrahydrofolate synthetase | Contains protein domain (PF01268) - Formate-tetrahydrofolate ligase | Synthase | 56182575, 22278996, 56994075, 264259, 29331822, 29331824, 29331826, 29331827, 29331828, 29146498, 29146499, 29331830, 265009, 60170831, 33657402, 33109954, 87168559, 265018, 18108351, 264448, 21906765, 21906767, 21906768, 29148627, 29148629, 29148784, 60170615, 52644150, 33657023, 33657109, 18108374, 55811576, 264559, 18108365, 22279000, 264563 |
| 1709 | 80222563 (3417, 3418) | | | UNCLASSIFIED | 264107, 55811957, 263974, 263976, 263977, 263981 |
| 1710 | 20754572 (3419, 3420) | Novel Protein sim. GBank gll5031735[ref NP_005760.1 pHEC - N-acetylglucosamine 6-O-sulfotransferase | | UNCLASSIFIED | 284556 |
| 1711 | 91013729 (3421, 3422) | Novel Protein sim. GBank gll5454168[ref NP_008453.1 pXAP4 - HBV associated factor | Contains protein domain (PF00641) - Zn-finger in Ran binding protein and others. | sulfotransferase | 65274572, 29331824, 29331826, 264768, 60431528, 35696423, 60432113, 264563 |
| 1712 | 95330184 (3423, 3424) | Novel Protein sim. GBank gll5454168[ref NP_008453.1 pXAP4 - HBV associated factor | Contains protein domain (PF00641) - Zn-finger in Ran binding protein and others. | kinase | 56994075, 264093, 264259, 29331822, 264099, 29331824, 29331827, 264107, 264110, 264511, 264592, 265011, 265018, 264683, 264686, 264689, 265020, 33657023, 263967, 33657109, 263974, 35696423, 35695855, 264630, 264636, 264558, 264566 |
| 1713 | 94143453 (3425, 3426) | Novel Protein sim. GBank gll160409 (M69183) - mature-parasite-infected erythrocyte surface antigen [Plasmodium falciparum] | Contains protein domain (PF00643) - B-box zinc finger. | UNCLASSIFIED | 22278985, 264508, 264758, 18108351, 18108370, 263974, 18108374, 264634, 56182323, 83373044, 60432113 |
| 1714 | 87420048 (3427, 3428) | | | | 22278997, 264757, 21906765, 265020, 265021, 264692, 56526486 |
| 1715 | 94260257 (3429, 3430) | Novel Protein sim. GBank gll5689537[dbj BAA83052.1] - (AB028023) KIAA1100 protein [Homo sapiens] | | UNCLASSIFIED | 264509, 264905, 264906, 264907, 264908, 264909, 264910, 264591, 265011, 264766, 264768, 264769, 264691, 264692, 264632, 264634, 264635, 264636, 264637, 264556, 264639, 264564 |
| 1716 | 87400449 (3431, 3432) | Novel Protein sim. GBank gll4589468[dbj BAA76761.1] - (AB012808) mBOCT [Mus musculus] | | transport | 56182575, 29331824, 60432289, 264109, 264909, 265007, 264600, 265019, 264886, 265020, 264693, 55811576, 264558, 60432113, 22279002 |

| | | | | | |
|------|-----------------------|---|--|--------------|--|
| 1717 | 87563223 (3433, 3434) | Novel Protein sim. GBank gij2765411[embjCAA74749] - (Y14391) GTP-binding protein [Homo sapiens] | | UNCLASSIFIED | 264569, 264259, 29331825, 29331826, 29331828, 35696052, 264508, 264905, 264907, 264908, 264909, 264512, 265009, 264910, 264592, 264758, 264759, 265017, 264681, 264764, 264766, 264688, 18108357, 35695917, 264690, 264692, 264693, 264628, 264629, 35696423, 264630, 264631, 264635, 264636, 18108380, 264638, 264639, 18108388, 18108391 |
| 1718 | 87032828 (3435, 3436) | Novel Protein sim. GBank gij2833262[spjQ14999Y076_HUMAN - HYPOTHETICAL PROTEIN KIAA0076 (HA0936)] | | UNCLASSIFIED | 265011, 264681, 264682, 264684, 264688, 264689, 21906765, 265021, 264691, 33657023, 264693, 18108370, 35695855, 264632, 264634, 264636, 18108388, 22279002 |
| 1719 | 94315259 (3437, 3438) | Novel Protein sim. GBank gij4505197[refjNP_003473.1]pMLL2 - myeloid/lymphoid or mixed-lineage leukemia 2 | | UNCLASSIFIED | 18108398, 65274572, 35696286, 22278997, 60432049, 56182181, 66714117, 60432289, 29331828, 35696052, 29331828, 264906, 29331830, 56182435, 264592, 60431735, 60433438, 55812038, 264759, 265010, 264600, 264601, 265017, 264448, 264764, 264288, 264769, 21906766, 21906769, 55811957, 265020, 265021, 52644150, 33657023, 33657109, 33657182, 27486262, 33657349, 35695763, 18108370, 60431528, 18108374, 35696423, 55811576, 35695855, 264631, 56182323, 264559, 264584, 264486, 56182575, 22278999, 264259, 29331824, 60432289, 29331827, 35696052, 264508, 264905, 264906, 264907, 264908, 264909, 264511, 264910, 264758, 21906754, 265011, 264601, 264760, 264762, 264288, 264766, 264886, 18108357, 264689, 21906765, 55811957, 264693, 20281149, 264629, 18108374, 55811576, 65274791, 264630, 20281071, 264634, 264635, 264638, 264637, 264556, 264638, 264639, 56182323, 87168518 |
| 1720 | 84853063 (3439, 3440) | Novel Protein sim. GBank gij2129478[pirj]IS51939 - chitinase (EC 3.2.1.14) precursor - beet | | UNCLASSIFIED | 22278994, 22278999, 29331822, 265006, 265007, 265008, 55812038, 21906754, 60174639, 265011, 87168559, 18108351, 18108354, 21906765, 21906766, 21906768, 21906769, 265020, 33657109, 18108370, 18108374, 264556, 60170394, 83373044, 18108385, 264486 |
| 1721 | 91722288 (3441, 3442) | Novel Protein sim. GBank gij4886461[embjCAB43381.1] - (AL050280) hypothetical protein [Homo sapiens] | | UNCLASSIFIED | 56994075, 29331824, 29331828, 265009, 18108351, 21906768, 265020, 33657023, 18108374, 83373044 |
| 1722 | 84134549 (3443, 3444) | Novel Protein sim. GBank gij5689375[dbj]BAA82968.1] - (AB030644) tudor repeat associator with PCTAIRE 2 [Rattus norvegicus] | Contains protein domain (PF00567) - Tudor domain | kinase | |

| | | | | | |
|------|-----------------------|---|--|--------------|--|
| 1723 | 95358181 (3445, 3446) | Novel Protein sim. GBank gi4426962[gb AAD20633] - (AF126062) Arf-like 2 binding protein BART1 [Homo sapiens] | | UNCLASSIFIED | 264488, 264687, 264769, 21906767, 21906768, 58182575, 55811957, 22278997, 22278998, 265020, 264259, 264692, 33657023, 29331822, 29331824, 29331825, 60432289, 33657182, 33856970, 33857349, 29146499, 264508, 264907, 18108370, 264628, 264908, 264909, 18108374, 55811576, 264510, 265008, 264511, 265007, 264910, 264632, 264591, 60432229, 264592, 60433356, 264594, 60433438, 264595, 83373044, 55812038, 33109954, 33657084, 87168518, 87168474, 265010, 265011, 87168559, 264600, 60432113, 264604, 265019, 264563, 264448, 264682, 264586, 264764, 264288, 264587, 264486, 264369, 264766 |
| 1724 | 87713806 (3447, 3448) | Novel Protein sim. GBank gi2340162 (AF005083) - dsRBP-ZFa [Xenopus laevis] | Contains protein domain (PF00096) - Zinc finger, C2H2 type | UNCLASSIFIED | 264905, 18108359, 264693, 264628, 264631, 264636, 264555, 264558, 264558, 264559, 35696286, 264259, 29331822, 35696052, 264508, 264509, 264905, 264906, 264907, 264908, 264909, 264910, 265009, 264591, 264601, 264760, 18108351, 264681, 264764, 264288, 264766, 264768, 21906769, 35695917, 264628, 35696423, 264630, 264631, 264632, 264635, 264636, 264638, 87168518, 264586 |
| 1725 | 85655191 (3449, 3450) | Novel Protein sim. GBank gi3152662 (AF064604) - KE03 protein [Homo sapiens] | Contains protein domain (PF00023) - homeobox Ank repeat | homeobox | |
| 1726 | 85754255 (3451, 3452) | | | UNCLASSIFIED | |
| 1727 | 85296382 (3453, 3454) | Novel Protein sim. GBank gi4689348[gb AAD27861.1 AF13256 - (AF132562) BcDNA.LD14270 [Drosophila melanogaster] | | UNCLASSIFIED | 29146498, 264683, 264689, 264905, 265011, 264689, 21906768 |

| | | | | |
|------|-----------------------|---|---|---|
| 1728 | 95349515 (3455, 3456) | Novel Protein sim. GBank gij4406549[gbjAAD20027] - (AF131738) Unknown [Homo sapiens] | UNCLASSIFIED | 60424179, 18108397, 56182575, 22278995, 56994075, 35696286, 22278997, 22278998, 22278999, 264094, 60432049, 264259, 29331822, 29331824, 56182181, 29331825, 60432289, 29331826, 29331827, 35696052, 264905, 264906, 264907, 29331830, 66712502, 264908, 56182435, 264511, 265008, 265009, 60432229, 60433356, 33657402, 60433438, 264759, 21908754, 87168474, 265010, 265011, 87168559, 265017, 265018, 265019, 5581150, 264681, 264448, 264682, 264763, 264683, 264288, 264684, 264368, 264685, 264766, 264687, 264769, 21908764, 264689, 21908765, 21908766, 21908767, 21908768, 35695917, 265020, 265021, 265022, 264535, 264691, 264692, 33657023, 264693, 33657109, 18108370, 264628, 263972, 264629, 18108374, 18108376, 55810764, 65274791, 35695855, 264631, 264634, 264635, 60431850, 264636, 264638, 60170394, 264639, 83373044, 56526486, 87168518, 60432113, 22279000, 22279002, 264564, 264566 |
| 1729 | 81227948 (3457, 3458) | Novel Protein sim. GBank gij854065[embjCAA58337] - (X83413) U88 [Human herpesvirus 6] | UNCLASSIFIED | 264906, 264907, 264908, 264511, 264555, 83373044, 264596, 264566 |
| 1730 | 85483474 (3459, 3460) | | UNCLASSIFIED | 29331822, 29331825, 29331828, 264907, 264908, 264909, 265011, 264764, 264629 |
| 1731 | 88266068 (3461, 3462) | Novel Protein sim. GBank gij831600[piljS47094 - hypothetical protein - rabbit] | UNCLASSIFIED | 5264842, 264907, 264909, 56182435, 55811386, 87168559, 265018, 265019, 264760, 52644229, 55811576 |
| 1732 | 81218878 (3463, 3464) | Novel Protein sim. GBank gij4240231[dbjBAA74894.1] - (AB020678) KIAA0871 protein [Homo sapiens] | struct | 56182575, 29331822, 29331824, 29331827, 66712502, 264591, 33657402, 60433356, 265019, 21908768, 21908769, 35695917, 265020, 265021, 264636, 56182323 |
| 1733 | 87617178 (3465, 3468) | Novel Protein sim. GBank gij1575756 (U70674) - m-Numb [Mus musculus] | Contains protein domain (PF00840) - synthase Phosphotyrosine interaction domain (PTB/PID). | 264907, 264910, 33657402, 265010, 264681, 264683, 264684, 264686, 264769, 264691, 264692, 264693, 264628, 264636, 264556, 264693 |
| 1734 | 87795261 (3467, 3468) | | | |

| | | | | |
|------|-----------------------|---|--|---|
| 1735 | 88318638 (3469, 3470) | Novel Protein sim. GBank gi4836807 gb AAD30566.1 AF146793 PF127 [Mus musculus] | Contains protein domain (PF01169) - Uncharacterized protein family UPF0016 | 264488, 18108394, 18108398, 22278998, 60432049, 264259, 29331822, 264908, 265006, 265007, 265008, 265009, 264591, 33657402, 265010, 265011, 87168559, 265017, 18108351, 264682, 18108354, 264769, 264689, 21906765, 21906768, 21906769, 29148629, 29148784, 265021, 265022, 52644150, 18108384, 18108385, 33657109, 18108370, 18108374, 18108380, 18108385, 87168518, 264563, 18108390, 264490, 264259, 66714117, 66712502. |
| 1738 | 95362884 (3471, 3472) | Novel Protein sim. GBank gi4885647 ref NP_005472.1 pTRAP - thyroid hormone receptor-associated protein complex component | UNCLASSIFIED | 56182435, 265006, 265008, 265009, 264910, 60433356, 87168559, 265017, 265019, 18108351, 265020, 265022, 33657023, 33657109, 264555, 60431850, 264637, 60170394, 264558, 264639, 264564 |
| 1737 | 88165549 (3473, 3474) | Novel Protein sim. GBank gi2143607 pir S68695 - BIK protein - rat | kinase | 29331822, 29331830, 264591, 265011, 265018, 265019, 22279002 |
| 1738 | 85788811 (3475, 3476) | Novel Protein sim. GBank gi2225941 emb CAA69714 - (Y08460) Mdes protein [Mus musculus] | UNCLASSIFIED | 264908, 264909, 265008, 264910, 264566 |
| 1739 | 87328576 (3477, 3478) | | UNCLASSIFIED | 35696052, 264603, 264557 |
| 1740 | 83592939 (3479, 3480) | Novel Protein sim. GBank gi4809 emb CAA44309 - (X62452) YCR601 [Saccharomyces cerevisiae] | traffic | 264604, 21906764, 18108364, 264629, 35695855, 264638 |
| 1741 | 95010100 (3481, 3482) | Novel Protein sim. GBank gi4883888 gb AAD31695.1 AF13042 - (AF130420) serine protease-like protein isoform [Homo sapiens] | UNCLASSIFIED | 60432289, 29331827, 264509, 265009, 60432229, 264759, 265017, 264767, 264688, 264689, 21906769, 265020, 33657109 |
| 1742 | 85788814 (3483, 3484) | Novel Protein sim. GBank gi4505183 ref NP_003667.1 pMLD - membrane fatty acid (lipid) desaturase | UNCLASSIFIED | 264908, 264910, 264758, 265011, 264631, 264638, 264566 |
| 1743 | 86966475 (3485, 3486) | | | 265017, 265020, 264692 |
| 1744 | 81224003 (3487, 3488) | Novel Protein sim. GBank gi728832 sp P39189 ALU2_HUMAN - III ALU SUBFAMILY SB WARNING ENTRY III | UNCLASSIFIED | 22278998, 264508, 264907 |
| 1745 | 20290075 (3489, 3490) | | | 264558 |
| 1746 | 84326110 (3491, 3492) | Novel Protein sim. GBank gi731756 sp P38873 YHY6_YEAST - HYPOTHETICAL 175.8 KD PROTEIN IN GND1-IK1 INTERGENIC REGION | UNCLASSIFIED | 65274791, 264639, 264559 |
| 1747 | 84324333 (3493, 3494) | Novel Protein sim. GBank gi1658503 (U75467) - Alu [Drosophila melanogaster] | transcription factor | 52646842, 29331824, 29331825, 52644045, 56182435, 265007, 52646317, 87168474, 265018, 18108351, 264369, 264769, 264689, 35695917, 60170615, 264691, 33657023, 65274791, 35695855, 60170394, 22279000, 264482 |

| | | | | | |
|------|-----------------------|--|---|--------------|--|
| 1748 | 88003580 (3495, 3498) | Novel Protein sim. GBank gi 4504511 ref NP_001530.1 phsJ2 - heat shock protein, DNAJ-like 2 | Contains protein domain (PF00684) - DnaJ central domain (4 repeats) | eph | 264489, 56182575, 29331824, 56182435, 264112, 265007, 265019, 264764, 21906768, 265020, 264691, 55811576, 264635, 264555, 264556, 264557, 264559 264106 |
| 1749 | 83363091 (3497, 3498) | Novel Protein sim. GBank gi 5650780 gb AAD45948.1 AF15196 (AF151968) RGS protein RGS-17 [Gallus gallus] | Contains protein domain (PF00615) - Regulator of G protein signaling domain | oncogene | |
| 1750 | 94321664 (3499, 3500) | Novel Protein sim. GBank gi 4996894 gb AAC28444.2 - (AF065164) hyperpolarization-activated, cyclic nucleotide- gated channel 2 [Homo sapiens] | | | 33657402, 264288, 52644150, 263974, 83373044 |
| 1751 | 83373058 (3501, 3502) | Novel Protein sim. GBank gi 2760161 gb BAA24184 - (AB010054) outer arm dynein light chain 2 [Anthriscus crassispina] | Contains protein domain (PF00560) - ATPase associated Leucine Rich Repeat | | 265010, 264369 |
| 1752 | 86456530 (3503, 3504) | Novel Protein sim. GBank gi 3915482 sp P74346 YG29_SYNY3 - HYPOTHETICAL 36.0 KD PROTEIN SLR1629 | Contains protein domain (PF00849) - RNA pseudouridylylase synthase | deaminase | 264510, 264593, 264682, 21906765, 18108370 |
| 1753 | 94235159 (3505, 3506) | Novel Protein sim. GBank gi 2852636 (AF007155) - unknown [Homo sapiens] | Contains protein domain (PF01553) - Acyltransferase | phosphatase | 56994075, 22278996, 264908, 60170831, 264682, 264764, 264369, 264288, 264685, 264687, 21906766, 264692, 264693, 65274620, 65274791, 35685855, 264637, 264564 |
| 1754 | 88095323 (3507, 3508) | Novel Protein sim. GBank gi 731421 sp P39981 YEH4_YEAST - HYPOTHETICAL 53.3 KD PROTEIN IN HXT8-CAN1 INTERGENIC REGION | | transport | 264488, 35696286, 264509, 264906, 264907, 264908, 264909, 264511, 264910, 264591, 33657402, 264594, 264757, 264758, 264600, 264604, 264762, 264763, 264685, 264766, 264691, 264628, 35696423, 264632, 264634, 264637, 264638, 263981, 264558, 264639, 264563, 264564, 264565, 264566, 264567 264686 |
| 1755 | 79470282 (3509, 3510) | Novel Protein sim. GBank gi 1176422 (U43194) - rhophilin [Mus musculus] | | UNCLASSIFIED | |
| 1756 | 92862614 (3511, 3512) | Novel Protein sim. GBank gi 4432860 gb AAD20708 - (AC006300) putative glucose-induced repressor protein [Arabidopsis thaliana] | | | 52846842, 22278994, 22278995, 56994075, 22278996, 22278997, 264259, 29331822, 60432289, 29331827, 33656970, 265006, 265009, 60432229, 60433356, 60433438, 33109954, 21908754, 265017, 265018, 265019, 264448, 264369, 264288, 21906765, 21908766, 21908767, 21908768, 21908769, 265020, 265021, 265022, 264692, 27486262, 27486264, 18108376, 20281152, 264558, 18108388, 87168518, 60432113, 22279000, 22279002, 264482 |
| 1757 | 95357380 (3513, 3514) | Novel Protein sim. GBank gi 5441615 emb CAB46856.1 - (AJ388557) zinc finger protein [Canis familiaris] | Contains protein domain (PF00096) - Zinc finger, C2H2 type | dna_rna_bind | 22278997, 264259, 60432289, 29331827, 264908, 52644045, 265008, 264593, 265019, 264766, 21908768, 65274620, 18108385, 60432113, 264566, 264487 |

| | | | | | |
|------|-----------------------|--|---|--------------|--|
| 1758 | 87612971 (3515, 3516) | Novel Protein sim. GBank gij3881040 emb CAA16403 - (AL021497) predicted using GeneFinder [Caenorhabditis elegans] | | UNCLASSIFIED | 22278994, 22278999, 29331822, 29331824, 29331825, 33656970, 264508, 265006, 265007, 265009, 264591, 33657402, 33109954, 87168474, 264600, 265017, 265018, 21906769, 265020, 265021, 33657023, 33657109, 264629, 18108374, 35695855, 264632, 52644332, 22278002, 264583 |
| 1759 | 36894372 (3517, 3518) | | | UNCLASSIFIED | 264759 |
| 1760 | 87328716 (3519, 3520) | Novel Protein sim. GBank gij3262749 emb CAB45688.1 - (AJ131320) Proline rich synapse associated protein 2 [Rattus norvegicus] | | UNCLASSIFIED | 56182575, 60432049, 35696052, 264805, 264906, 264907, 264908, 264909, 265006, 265009, 264910, 60432229, 264592, 264595, 55812038, 264758, 264762, 18108351, 264764, 264766, 264768, 264769, 21806765, 55811957, 35695917, 264690, 264892, 264628, 264629, 55811576, 35698423, 264632, 264634, 264636, 264557, 264639, 60432113, 264565, 264486 |
| 1761 | 87409586 (3521, 3522) | Novel Protein sim. GBank gij127749 sp P10569 MYSC_ACACA - MYOSIN IC HEAVY CHAIN | Contains protein domain (PF00618) - Guanine nucleotide exchange factor for Ras-like GTPases; N-terminal motif | UNCLASSIFIED | 29331822, 264910, 264685, 264686 |
| 1762 | 95319887 (3523, 3524) | Novel Protein sim. GBank gij3169158 (AC004770) - BC289730_2 [Homo sapiens] | Contains protein domain (PF00173) - Heme-binding domain in cytochrome b5 and oxidoreductases | desaturase | 22278999, 264259, 264905, 264907, 60170831, 265010, 265011, 265017, 264448, 21906765, 21906766, 21906767, 21906768, 265021, 264690, 33657109, 18108374, 264558, 60170394 |
| 1763 | 91224013 (3525, 3526) | Novel Protein sim. GBank gij4809026 gb AAD30062.1 - (AF132856) suppressor of G2 allele of skp1 homolog [Homo sapiens] | | | 56181686, 29331825, 35696052, 264905, 264906, 264909, 264763, 264682, 264769, 35695917, 265022, 33657023, 18108374, 35696423, 264634 |
| 1764 | 87757687 (3527, 3528) | Novel Protein sim. GBank gij1360869 prj CGHUV - collagen alpha 1(V) chain precursor - human | Contains protein domain (PF01410) - Fibrillar collagen C-terminal domain | collagen | 22278998, 264259, 35696052, 29331828, 56182435, 265008, 265017, 265018, 264448, 264288, 21906766, 21906767, 29148627, 35695917, 264691, 33657023, 60432113, 22279002 |
| 1765 | 81230091 (3528, 3530) | Novel Protein sim. GBank gij486806 prj JS35503 - finger protein neutralized - fruit fly (Drosophila melanogaster) | | UNCLASSIFIED | 264488, 18108394, 22278995, 22278997, 22278999, 264259, 264104, 264508, 264905, 264907, 264511, 265007, 265008, 264910, 265009, 21906754, 265010, 265017, 264603, 265018, 265019, 18108351, 264682, 264448, 264369, 264288, 264766, 18108359, 21906766, 21906767, 29148627, 29148629, 35695917, 265020, 265021, 264692, 264628, 264629, 18108374, 263976, 264636, 83373044, 22279000, 22279002 |

| | | | | | |
|------|-----------------------|--|---|-------------------|--|
| 1766 | 95081201 (3531, 3532) | Novel Protein sim. GBank gi2499087 sp Q09332 UGG_DROME - UDP- GLUCOSE-GLYCOPROTEIN GLUCOSYLTRANSFERASE PRECURSOR (DUGT) | | glycoprotein | 52645156, 87168559, 60170615, 33657023, 264693, 33657109, 27486261, 264555, 83373044 |
| 1767 | 87755998 (3533, 3534) | Novel Protein sim. GBank gi1152042 sp P00736 C1R_HUMAN - COMPLEMENT C1R COMPONENT PRECURSOR | | UNCLASSIFIED | 265017, 265019, 264686, 264768, 265020, 264692 |
| 1768 | 80253216 (3535, 3536) | | | UNCLASSIFIED | 29331824, 29331825, 264591, 56182323 |
| 1769 | 87388988 (3537, 3538) | | | UNCLASSIFIED | 264563 |
| 1770 | 95413144 (3539, 3540) | Novel Protein sim. GBank gi1152042 sp P00736 C1R_HUMAN - COMPLEMENT C1R COMPONENT PRECURSOR | Contains protein domain (PF00089) - Trypsin | complement | 264488, 264768, 264769, 56182575, 55811957, 264690, 264691, 35696052, 264905, 264509, 264906, 264907, 264628, 264908, 264909, 264910, 264634, 264635, 264636, 264556, 264757, 264758, 55812038, 65274444, 264760, 264563, 264762, 264764, 264684, 264766 |
| 1771 | 94233542 (3541, 3542) | Novel Protein sim. GBank gi3914191 sp P56558 OGT1_RAT - UDP-N- ACETYLGLUCOSAMINE-PEPTIDE N- ACETYLGLUCOSAMINYLTRANSFERASE 110 KD SUBUNIT (O-GLCNAC TRANSFERASE P110 SUBUNIT) | Contains protein domain (PF00515) - TPR Domain | transferase | 264758, 264600, 264369, 55811957, 265020, 83373044, 22278000 |
| 1772 | 87643510 (3543, 3544) | Novel Protein sim. GBank gi4959442 gb AAD34351.1 AF12136 - (AF121360) DNZDHCNEW1 zinc finger protein 11 [Drosophila melanogaster] | Contains protein domain (PF01529) - DHC zinc finger domain | | 22278998, 29331828, 33109954, 265018, 265019, 264764, 21906765, 265020, 265021, 264556 |
| 1773 | 94116824 (3545, 3546) | Novel Protein sim. GBank gi3978464 (AF085693) - G protein-coupled receptor kinase-associated ADP ribosylation factor GTPase-activating protein [Rattus norvegicus] | Contains protein domain (PF01412) - Putative GTP-ase activating protein for Arf | potassium_channel | 65274572, 56182575, 22278996, 35696286, 22278999, 264259, 29331824, 60424269, 29331825, 60432289, 35696052, 264106, 264509, 264906, 264907, 29331830, 264908, 52644045, 264511, 265006, 265007, 265008, 60170831, 60433438, 264758, 55811388, 87168559, 265017, 264604, 265019, 55811150, 264288, 56181562, 264689, 21906768, 21906767, 21906768, 21906769, 55811957, 265020, 265022, 52844150, 264691, 33657023, 264692, 264693, 60431528, 35696423, 35695855, 264636, 56182323, 18108387, 56526486, 22278000, 22279002, 264563, 264564, 264565, 264566, 264567 |
| 1774 | 94232573 (3547, 3548) | Novel Protein sim. GBank gi2495699 sp Q15034 Y032_HUMAN - HYPOTHETICAL PROTEIN KIAA0032 | Contains protein domain (PF00415) - Regulator of chromosome condensation (RCC1) | UNCLASSIFIED | 65274572, 56182575, 35696052, 55812038, 33109954, 21908754, 265017, 21908767, 21906768, 21906769, 265020, 264691, 264636, 56182323, 22279002 |

| | | | | | |
|------|-----------------------|--|--------------|---|--|
| 1775 | 95359330 (3549, 3550) | Novel Protein sim. GBank gij1469199[dbj BAA09487] - (D50928) The KIAA0138 gene product is novel. [Homo sapiens] | | UNCLASSIFIED | 85274572, 56182575, 56994075, 35696286, 22278997, 29331822, 29331826, 60432289, 29331828, 35696052, 29331830, 66712502, 264828, 56182435, 264511, 265007, 265009, 60170831, 60432229, 60433438, 55812038, 21906754, 85658542, 87168559, 264601, 265017, 265018, 265019, 264762, 264448, 264288, 264889, 21906765, 21906766, 21906767, 21906768, 265020, 265021, 265022, 264691, 33657023, 264692, 33657109, 27486261, 33657349, 18108370, 18108377, 35696423, 55811576, 35695855, 264632, 264634, 264636, 264639, 56182323, 83373044, 56526486, 87168518, 60432113, 22279000, 22279002, 264482, 264486, 264910 |
| 1776 | 94133756 (3551, 3552) | Novel Protein sim. GBank gij4589676[dbj BAA76857.1] - (AB023230) KIAA1013 protein [Homo sapiens] | | | |
| 1777 | 87447171 (3553, 3554) | Novel Protein sim. GBank gij3219939[sp P87115 YDK9_SCHPO - HYPOTHETICAL 116.5 KD PROTEIN C20G8.09C IN CHROMOSOME 1 | nucl_rept | | 56994075, 29331826, 265008, 87168474, 265017, 265018, 264761, 55811150, 264764, 56181562, 264689, 21906765, 21906768, 21906769, 35695917, 264690, 33657023, 35695763, 60431528, 35696423, 55811576, 35695855, 22279000, 22279002, 264564 |
| 1778 | 94851624 (3555, 3556) | Novel Protein sim. GBank gij3875848[emb CAA91454.1] - (Z66561) Similarity to Human rab13 protein (PIR Acc. No. A49847). Contains the ATP/GTP-binding site motif (PROSITE PS00017); cDNA EST EMBL.M89412 comes from this gene; cDNA EST yk212g9.3 comes from this gene; cDNA EST yk212g9... | UNCLASSIFIED | | 29331826, 29331827, 35696052, 264512, 265007, 265009, 265017, 265019, 264762, 18108351, 264769, 21906765, 21906766, 21906767, 21906768, 21906769, 265020, 265021, 264691, 264693, 18108370, 264556, 83373044 |
| 1779 | 94133758 (3557, 3558) | Novel Protein sim. GBank gij4589676[dbj BAA76857.1] - (AB023230) KIAA1013 protein [Homo sapiens] | UNCLASSIFIED | | 29148827, 35696286, 29147620, 265006, 265007, 265008, 18108385, 65274727, 264482, 264369, 264766 |
| 1780 | 87023497 (3559, 3560) | | UNCLASSIFIED | Contains protein domain (PF00807) - Apidaecin | 264107, 33657109, 56526486 |
| 1781 | 84047477 (3561, 3562) | | UNCLASSIFIED | | 264508, 264906, 264639 |
| 1782 | 88094607 (3563, 3564) | Novel Protein sim. GBank gij729225[sp P41237 CTXN_RAT - CORTEXIN | | | 264259, 29331822, 264508, 264905, 264906, 264907, 264908, 265007, 265009, 264910, 264591, 264758, 264764, 264288, 264768, 264769, 264635, 264636, 264637, 264639, 264553 |
| 1783 | 85717905 (3565, 3566) | Novel Protein sim. GBank gij2257543[dbj BAA21436] - (AB004538) protein arginine N-methyltransferase [Schizosaccharomyces pombe] | interferon | | 264768 |

| | | | | |
|------|-----------------------|---|--|--|
| 1784 | 95187093 (3567, 3568) | Novel Protein sim. GBank gi 1755049 (U55042) - myosin X (Bos taurus) | Contains protein domain (PF00169) - struct PH domain | 35696286, 264259, 35696052, 264508, 264905, 264906, 264907, 66712502, 264908, 264909, 265007, 265008, 265009, 264910, 264591, 264594, 264757, 264758, 264759, 265010, 265011, 264601, 264602, 264604, 264605, 18108351, 264762, 264763, 264764, 264389, 264766, 264687, 264768, 264688, 21908768, 35695917, 264690, 264691, 264692, 264693, 264628, 18108374, 35696423, 264631, 264632, 264635, 264637, 264638, 264639, 18108385, 22279000, 22279002, 264565, 264566, 264488 |
| 1785 | 95357475 (3569, 3570) | Novel Protein sim. GBank gi 4589552 dbj BAA76798.1 - (AB023171) KIAA0954 protein [Homo sapiens] | UNCLASSIFIED | 65274572, 264259, 29331822, 29331824, 29331825, 29331827, 29331828, 264905, 264906, 264908, 66712502, 56182435, 264511, 265007, 60433356, 55811150, 264683, 264369, 264687, 52644229, 21908767, 52644150, 33657023, 65274620, 33657182, 65274791, 35695855, 264555, 65274727, 22279002 |
| 1786 | 85296485 (3571, 3572) | Novel Protein sim. GBank gi 117788 sp P26770 CYAA_RAT ADENYLATE CYCLASE, TYPE IV (ATP PYROPHOSPHATE-LYASE) (ADENYLYL CYCLASE) | UNCLASSIFIED | 264908, 35696423, 264636 |
| 1787 | 87434784 (3573, 3574) | Novel Protein sim. GBank gi 3877175 emb CAA90338.1 - (Z50028) cDNA EST vk321h8.5 comes from this gene; cDNA EST EMBL:D68896 comes from this gene; cDNA EST vk395f9.5 comes from this gene [Caenorhabditis elegans] | UNCLASSIFIED | 264488, 264905, 264908, 264909, 264595, 264764, 264766, 264692, 60431528, 264629, 264636, 264564, 264566 |
| 1788 | 91228779 (3575, 3576) | | UNCLASSIFIED | 264488, 83373044 |
| 1789 | 88094529 (3577, 3578) | Novel Protein sim. GBank gi 2088669 (AF003130) - F55A12.9 gene product [Caenorhabditis elegans] | UNCLASSIFIED | 264488, 29331828, 264909, 18108351, 264288, 265021, 264555, 264636 |
| 1790 | 82489734 (3579, 3580) | | | 35696052, 264905, 264906, 264907, 264908, 264909, 265008, 264910, 264758, 265011, 265019, 264764, 264766, 264769, 264628, 264635 |

| | | | | | |
|------|-----------------------|--|--|-------------------|--|
| 1791 | 95197259 (3581, 3582) | Novel Protein sim. GBank gij2114321dbj BAA200371 - (D88733) membrane glycoprotein [Equine herpesvirus 1] | Contains protein domain (PF00047) - Immunoglobulin domain | glycoprotein | 264488, 264686, 264687, 264768, 18108394, 264769, 18108397, 264259, 264691, 264692, 33657023, 264693, 264509, 264905, 264906, 264628, 264907, 264629, 264908, 264909, 264510, 265008, 264511, 265008, 264630, 265009, 264631, 264910, 264632, 264634, 264635, 264555, 264636, 264592, 264637, 264593, 264638, 18108381, 264639, 264758, 265010, 265011, 264602, 22279000, 264604, 264760, 264564, 264681, 264762, 264565, 264763, 264683, 264586, 264764, 264288, 264684, 264567, 18108354, 18108391, 264685, 264766 |
| 1792 | 87792690 (3583, 3584) | Novel Protein sim. GBank gij4337106 gb AAD180821 - (AF129756) BAT4 [Homo sapiens] | Contains protein domain (PF01585) - G-patch domain | UNCLASSIFIED | 22278997, 264259, 264508, 265007, 33657402, 87168559, 264369, 33657023, 35695855, 20281071, 264559, 18108387, 87168518 |
| 1793 | 95337877 (3585, 3586) | Novel Protein sim. GBank gij5579331 gb AAD45504.1 AF145732 - (AF145732) endoplasmic reticulum alpha-mannosidase I [Homo sapiens] | Contains protein domain (PF01532) - Glycosyl hydrolase family 47 | ATPase-associated | 65274572, 22278995, 22278996, 22278997, 22278998, 264093, 264259, 29331824, 66714117, 60432289, 29331827, 29331828, 264103, 264105, 29331830, 265007, 264910, 265009, 60170831, 60433356, 21906754, 265010, 265017, 265019, 264681, 264682, 264288, 52644229, 21906765, 21906766, 21908767, 21906768, 21906769, 265020, 265021, 265022, 60170615, 52644150, 33657023, 33657109, 18108370, 18108374, 65274791, 20281071, 60432113, 22278000, 264482, 264564 |
| 1794 | 87759808 (3587, 3588) | Novel Protein sim. GBank gij4914604 emb CAB43677.1 - (AL050369) hypothetical protein [Homo sapiens] | Contains protein domain (PF01788) - Putative snoRNA binding domain | UNCLASSIFIED | 18108394, 22278995, 22278999, 264259, 29331822, 29331824, 29331825, 29146498, 29146499, 264508, 264905, 52644045, 264112, 265006, 265008, 264910, 60433356, 264757, 55812038, 87168474, 265011, 265017, 18108351, 264763, 264448, 264683, 264369, 21906765, 21906766, 21906767, 21906769, 29148784, 35695917, 60170815, 33657023, 264629, 18108374, 18108376, 35696423, 35695855, 264556, 264557, 264638, 264558, 18108385, 264564, 264632, 264635, 264636, 264595, 264596, 264907, 264566, 264909 |
| 1795 | 79747856 (3589, 3590) | | | UNCLASSIFIED | 264488, 264907, 264909, 264594, 264595, 264766, 264687, 21906765, 21906767, 264628, 264630, 264559 |
| 1796 | 86598486 (3591, 3592) | Novel Protein sim. GBank gij585084 sp Q07803 EFGM_RAT - ELONGATION FACTOR G, MITOCHONDRIAL PRECURSOR (MEF-G) | | glycoprotein | |

| | | | | | |
|------|-----------------------|---|--|---------------|--|
| 1797 | 91223219 (3593, 3594) | Novel Protein sim. GBank gii1842111 (U87586) - decoy (Arabidopsis thaliana) | | ribosomalprot | 22278996, 22278997, 22278998, 22278999, 29331822, 264910, 60170831, 21906754, 52644229, 21906765, 21906768, 21906769, 35695917, 265022, 52644150, 264691, 33657023, 263967, 33657109, 22279000 |
| 1798 | 91221276 (3595, 3596) | Novel Protein sim. GBank gii2832806(dbi)BAA24508.1) - (D89340) dipeptidyl peptidase III [Rattus norvegicus] | | peptidase | 22278994, 56994075, 22278997, 22278998, 22278999, 264259, 29331826, 60432289, 29331828, 33656970, 265008, 60432229, 264757, 60433438, 21906754, 33657084, 87168559, 265017, 18108351, 264682, 264448, 264288, 21906765, 21906768, 21906767, 21906768, 21906769, 35695917, 265020, 265021, 33657023, 33657182, 27486261, 27486265, 33657349, 263973, 18108374, 55811576, 35695855, 18108385, 87168518, 22279000, 264486 |
| 1799 | 86321713 (3597, 3598) | Novel Protein sim. GBank gii5689541(dbi)BAA83034.1) - (AB029025) KIAA1102 protein [Homo sapiens] | | eph | 264908, 21906754, 21906767, 21906769, 265020, 33657023, 264692, 264693, 264404, 22279000 |
| 1800 | 87080116 (3599, 3600) | Novel Protein sim. GBank gii4680879(gbi)AAD27729.1(AF13295) CGI-20 protein [Homo sapiens] | | UNCLASSIFIED | 264691, 264558, 264566 |
| 1801 | 95060723 (3601, 3602) | Novel Protein sim. GBank gii134920(sp)P21997(SSGP VOLCA - SULFATED SURFACE GLYCOPROTEIN 185 (SSG 185) | | | 52644045, 265007, 264632 |
| 1802 | 8771012 (3603, 3604) | Novel Protein sim. GBank gii134920(sp)P21997(SSGP VOLCA - SULFATED SURFACE GLYCOPROTEIN 185 (SSG 185) | | | 35696286, 66714117, 264508, 264509, 58182435, 264512, 18108351, 264688, 55811957, 264692, 55811576, 35695855, 264486 |
| 1803 | 95060725 (3605, 3606) | Novel Protein sim. GBank gii4680879(gbi)AAD27729.1(AF13295) CGI-20 protein [Homo sapiens] | | | 264686, 264488, 264687, 264489, 264768, 264769, 264689, 21906769, 35696286, 35695917, 264259, 264691, 264692, 264693, 20281099, 18108364, 35696052, 264508, 264509, 264905, 264906, 18108370, 264628, 264907, 66712502, 264908, 264909, 18108374, 18108376, 35696423, 35695855, 264510, 264511, 265008, 265007, 264512, 265008, 264910, 264631, 264632, 264634, 264635, 264591, 264638, 264637, 264592, 264638, 264593, 264639, 264594, 83373044, 264758, 264596, 18108385, 18108387, 265011, 264760, 264563, 18108351, 264762, 264564, 264448, 264585, 264763, 264683, 264764, 264566, 264288, 264486, 264567, 264765, 264766 |

| | | | | | |
|------|-----------------------|---|-----------------------|-------------------------------------|--|
| 1804 | 87770203 (3607, 3608) | Novel Protein sim. GBank gij3879914[embjCAA98538.1]- (Z74043) predicted using GeneFinder: cDNA EST EMBL:C13850 comes from this gene; cDNA EST EMBL:C11575 comes from this gene; cDNA EST yk3434.5 comes from this gene [Caenorhabditis elegans] | | | 52646365, 22278997, 22278999, 264905, 264908, 264909, 264910, 21906754, 264766, 21906765, 21906768, 35695917, 265020, 265022, 264691, 264637, 264639, 22279000, 264564, 264566 |
| 1805 | 95330375 (3609, 3610) | Novel Protein sim. GBank gij5453644[refNIP_006461.1]pEBBP - estrogen-responsive B box protein | | | 29331824, 29331825, 29331826, 29331827, 29331828, 87168559, 264288, 264687, 52644229, 35696423, 264636, 60432113 |
| 1806 | 94133762 (3611, 3612) | Novel Protein sim. GBank gij4589676[idj]BAA76857.1- (AB023230) KIAA1013 protein [Homo sapiens] | struct | | 264094, 264105, 264908, 35696423, 265006, 265007, 265008, 264555, 264592, 265011, 265016, 264369 |
| 1807 | 86943032 (3613, 3614) | | | | 29331824, 264908, 264910, 33657023, 263978 |
| 1808 | 87642711 (3615, 3616) | Novel Protein sim. GBank gij4884079[embjCAB43235.1]- (AL050008) hypothetical protein [Homo sapiens] | UNCLASSIFIED | | 264488, 35696286, 66714117, 35696052, 66712502, 264592, 60433438, 52644296, 265010, 264683, 264369, 264689, 55811957, 35695917, 33657109, 35695763, 55810764, 18108379, 35696423, 35695855, 56182323, 264563, 264564, 264487 |
| 1809 | 95321468 (3617, 3618) | Novel Protein sim. GBank gij1916927 (U87965) - putative G protein [Mus musculus] | UNCLASSIFIED | | 284594, 55811150, 264686, 29148629, 29148784, 264690, 264629, 18108374, 264556, 264557, 264558 |
| 1810 | 88096316 (3619, 3620) | Novel Protein sim. GBank gij1352944[spjP47178]YJSP_YEAST - HYPOTHETICAL 118.4 KD PROTEIN IN BAT2-DAL5 INTERGENIC REGION PRECURSOR | UNCLASSIFIED | | 264488, 35696052, 264905, 264906, 264907, 264908, 264909, 264511, 265009, 264910, 264592, 264593, 264594, 33657402, 264757, 264595, 264758, 264596, 264759, 264600, 264601, 264762, 264683, 264764, 264288, 264684, 264766, 264767, 264686, 264768, 264687, 264769, 264689, 265021, 264690, 264691, 264693, 264628, 264629, 18108374, 264630, 264631, 264632, 264634, 264635, 264636, 264637, 264638, 264639, 264563, 264566, 264486, 264567 |
| 1811 | 88086272 (3621, 3622) | Novel Protein sim. GBank gij2134984[pirj]J37275 - death-associated protein kinase (EC 2.7.1.-) - human | kinase Antk repeat | Contains protein domain (PF00023) - | 264488, 264759, 264508, 264509, 264905, 264906, 264907, 56182435, 264511, 264512, 264910, 264758, 265011, 264600, 264604, 18108354, 264766, 264686, 264769, 264534, 60170615, 33657023, 264629, 264631, 264639, 264563, 264482, 264483 |
| 1812 | 79245772 (3623, 3624) | | | | 29331822, 29331824, 265019, 18108351, 21906769 |

| | | | | | |
|------|-----------------------|--|---|------------------|---|
| 1813 | 88090972 (3625, 3626) | Novel Protein sim. GBank gi 5051636 gb AAD38326.1 AF07372 - (AF073727) EH domain-binding mitotic phosphoprotein [Homo sapiens] | Contains protein domain (PF01417) - ENTH domain | glucoamylase | 56182575, 264259, 29331824, 66714117, 29331828, 35696052, 264509, 264905, 264906, 264907, 264908, 66712502, 264909, 265007, 264910, 264591, 264593, 55812038, 265011, 265018, 264760, 264682, 264764, 264683, 264369, 264766, 264768, 264769, 21906768, 21906768, 264691, 264693, 18108374, 35695855, 264634, 264635, 264637, 264639, 264559, 22278000, 22278002, 264566 |
| 1814 | 88178047 (3627, 3628) | Novel Protein sim. GBank gi 3843608 (AC005395) - hypothetical protein [Arabidopsis thaliana] | | UNCLASSIFIED | 264488, 35696286, 22278998, 264092, 264094, 264259, 29331822, 29331824, 29331825, 60432289, 29331826, 29331827, 29331828, 264105, 264107, 52644045, 56182435, 265009, 60432229, 60433356, 87168474, 87168559, 264369, 264288, 21906765, 35695917, 265021, 265022, 33657023, 33657109, 18108374, 35696423, 264638, 55528486, 264482 |
| 1815 | 85296473 (3629, 3630) | Novel Protein sim. GBank gi 117788 sp P26770 CYA4_RAT ADENYLATE CYCLASE, TYPE IV (ATP PYROPHOSPHATE-LYASE) (ADENYL CYCLASE) | | siuct | 22278999, 264508, 264509, 264907, 264908, 264910, 265011, 264760, 264766, 264634, 264636 |
| 1816 | 83738845 (3631, 3632) | Novel Protein sim. GBank gi 1178623 sp P41846 YO96_CAEEL - HYPOTHETICAL 93.9 KD PROTEIN T20B12.6 IN CHROMOSOME III | | UNCLASSIFIED | 18108394, 18108397, 264509, 264907, 264908, 264909, 265009, 264591, 265011, 265017, 264687, 264688, 265022, 264691, 18108362, 18108368, 18108370, 18108374, 18108379, 264635, 264557, 264564, 264567, 264488, 22278997, 22278999, 60432049, 29331822, 29331824, 60432289, 52644045, 60170831, 265017, 265018, 265019, 18108351, 264682, 52644229, 21906765, 21906767, 21906768, 52644150, 33657023, 33657109, 27486262, 18108370, 18108374, 60170394, 56182323, 22279002 |
| 1817 | 88095268 (3633, 3634) | Novel Protein sim. GBank gi 3766377 emb CAA21429 - (AL031907) hypothetical protein [Schizosaccharomyces pombe] | Contains protein domain (PF00400) - WD domain, G-beta repeat | histone | |
| 1818 | 85806775 (3635, 3636) | Novel Protein sim. GBank gi 3879121 emb CAA94370 - (Z70310) predicted using GeneFinder; Similarity to Mouse ankyrin (PIR Acc. No. S37771); cDNA EST EMBL: T01923 comes from this gene; cDNA EST EMBL: D32335 comes from this gene; cDNA EST EMBL: D32723 comes from this gene; cDNA EST E... | Contains protein domain (PF00023) - Ank repeat | transcriptfactor | 35696286, 60433356, 264758, 264369, 264686, 21906769, 264693, 264632 |
| 1819 | 87759572 (3637, 3638) | Novel Protein sim. GBank gi 5031865 ref NP_005771.1 pLHFP - lipoma HMGIC fusion partner | | UNCLASSIFIED | 22278995, 22278997, 22278999, 264259, 29331826, 264908, 265007, 265008, 265009, 60432229, 265017, 265018, 265019, 264448, 264288, 21906768, 21906769, 265020, 18108381, 18108384, 22279000, 22279002, 264567 |

| | | | | | |
|------|-----------------------|--|---|-------------------|--|
| 1820 | 87769455 (3839, 3840) | | | | 284905, 284907, 284594 |
| 1821 | 80431510 (3841, 3842) | | | | 284907, 284768, 263978 |
| 1822 | 91221523 (3843, 3844) | Novel Protein sim. GBank gii4884130[emb]CAB43272.1] - (AL050101) hypothetical protein [Homo sapiens] | | | 22278995, 56994075, 22278996, 22278997, 22278998, 264259, 29331824, 29331825, 29331826, 35698052, 29331828, 264808, 29331830, 60170831, 264591, 264593, 60433356, 264598, 265017, 265019, 18108351, 264763, 264683, 21906765, 21906767, 21906768, 21906769, 35695917, 265020, 265021, 33657023, 18108364, 18108370, 35695855, 22278000, 22279002 |
| 1823 | 85522330 (3845, 3846) | | | UNCLASSIFIED | 264488, 264259, 264511, 264288, 264768, 264693, 35698423, 284634, 18108385, 264486 |
| 1824 | 86612025 (3847, 3848) | Novel Protein sim. GBank gii477072[pr]IA48018 - mucin 7 precursor, salivary - human | Contains protein domain (PF00047) - Immunoglobulin domain | UNCLASSIFIED | 284907, 264908, 264909, 284511, 264631, 264634, 264635, 264637, 264638, 264639, 264758, 264568 |
| 1825 | 87430125 (3849, 3850) | Novel Protein sim. GBank gii3036803[emb]CAA18493] - (AL022373) hypothetical protein [Arabidopsis thaliana] | | UNCLASSIFIED | 60432049, 264910, 264487 |
| 1826 | 91723612 (3851, 3852) | Novel Protein sim. GBank gii4680885[gb]AAD27732.1]AF13295 - (AF132957) CGI-23 protein [Homo sapiens] | | ATPase-associated | 52644507, 52645156, 52646842, 22278994, 22278996, 56994075, 264259, 60432049, 52645080, 35698052, 66712502, 52644045, 265008, 265009, 60432229, 60433356, 60433438, 52646317, 52644296, 265011, 87168559, 264448, 264288, 264369, 264688, 52644229, 264689, 21906765, 21906768, 265020, 60170615, 52644150, 33657023, 27486262, 27486264, 27486265, 35695763, 35698423, 35695855, 83373044, 87168518, 264404, 22279002 |
| 1827 | 81647212 (3853, 3854) | | | | 264758 |

| | | | | | |
|------|-----------------------|---|--|--------------|--|
| 1828 | 95074017 (3655, 3656) | Novel Protein sim. GBank gi 4503571 ref NP_001419.1 pENO1 - endolase 1, (alpha) | Contains protein domain (PF00113) - Enolase | oncogene | 284488, 5264642, 56182575, 22278996, 35696286, 22278997, 22278999, 264091, 264093, 60432049, 264259, 29331822, 29331824, 66714117, 29331825, 60432289, 29331826, 29331827, 29331828, 264105, 264508, 264907, 66712502, 52644045, 56182435, 265008, 264511, 264512, 265007, 265008, 265009, 60170831, 60432229, 264593, 60433356, 60433438, 264758, 33109954, 21906754, 87169474, 265010, 265011, 87168559, 265017, 265019, 264761, 264762, 264448, 264764, 264683, 264288, 264369, 18108355, 264768, 18108357, 18108358, 264688, 264769, 264689, 21906768, 21906769, 35695917, 265021, 60170615, 33657023, 33657349, 263972, 55811578, 35695856, 264635, 264555, 264558, 264638, 264557, 87168518, 22279000, 22279002, 264563, 264482, 264565, 264484, 264567 |
| 1829 | 80197720 (3657, 3658) | | | | 264508, 264634, 284509, 264482, 29331827, 264908, 265009, 264910 |
| 1830 | 94312942 (3659, 3660) | Novel Protein sim. GBank gi 2246532 (U93872) - ORF 73, contains large complex repeat CR 73 [Kaposi's sarcoma- associated herpesvirus] | | nuclease | 52645156, 22278994, 22278995, 35696286, 22278996, 22278997, 22278998, 22278999, 29331822, 29331825, 35696052, 52646317, 52644296, 87168559, 265019, 21906765, 21906766, 21906767, 21906768, 21906769, 35695917, 265021, 33657023, 52645129, 33657109, 33657182, 27486261, 27486262, 35695763, 263974, 35698423, 35695855, 52644332 |
| 1831 | 94138063 (3661, 3662) | | | UNCLASSIFIED | 29331824, 35696052, 29331830, 264595, 264758, 265010, 265019, 265022, 264693, 65274791 |
| 1832 | 84521663 (3663, 3664) | Novel Protein sim. GBank gi 1330345 (U58755) - coded for by C. elegans cDNA yk34b1.5; coded for by C. elegans cDNA yk13h10.5; coded for by C. elegans cDNA yk46e8.5; coded for by C. elegans cDNA yk46d5.5; coded for by C. elegans cDNA yk43c2.5; coded for by C. elegans cDNA yk46e8.... | | | 264602 |

| | | | | | |
|------|-----------------------|--|--|--------------|---|
| 1833 | 85314184 (3665, 3666) | Novel Protein sim. GBank gi 5174413 ref NP_006028.1 pCDC4 - CDC42-binding protein kinase beta (DMPK-like) | Contains protein domain (PF00069) - Eukaryotic protein kinase domain | kinase | 22278994, 22278997, 22278998, 22278999, 264259, 29331822, 29331824, 60432289, 29331827, 35686052, 29148499, 264508, 264509, 264808, 264907, 66712502, 264908, 52844045, 264909, 264512, 265008, 264591, 264593, 60433356, 21906754, 33657084, 265011, 265017, 264604, 265018, 265019, 264681, 18108351, 264683, 264288, 264685, 264766, 264687, 21906765, 21906766, 21906767, 21906768, 21906769, 29148629, 265020, 265021, 264690, 264692, 33657023, 65274620, 33657182, 27486284, 33657349, 65274791, 264634, 264635, 264558, 264557, 264558, 264559, 18108385, 56528486, 87168518, 60432113, 22279000, 22279002, 264563 |
| 1834 | 80562760 (3667, 3668) | | | | 264259, 264907, 264689, 22279000, 22279002 |
| 1835 | 94135718 (3669, 3670) | | | UNCLASSIFIED | 22278998, 29331822, 29331826, 87168474, 264603, 21906768, 263976, 35685855, 83373044 |
| 1836 | 87348450 (3671, 3672) | Novel Protein sim. GBank gi 4759286 ref NP_004268.1 pUCP4 - uncoupling protein 4 | Contains protein domain (PF00153) - Mitochondrial carrier proteins | transport | 29331825, 264908, 265019, 264764, 264686, 21906765, 264635 |
| 1837 | 94234287 (3673, 3674) | Novel Protein sim. GBank gi 3334400 sp Q24574 UBPE_DROME - UBIQUITIN CARBOXYL-TERMINAL HYDROLASE 64E (UBIQUITIN THIOLESTERASE 64E) (UBIQUITIN-SPECIFIC PROCESSING PROTEASE 64E) (DEUBIQUITINATING ENZYME 64E) | Contains protein domain (PF00443) - Ubiquitin carboxyl-terminal hydrolase family 2 | ubiquitin | 22278995, 29148499, 265006, 265008, 265009, 265010, 264683, 21906765, 29148627, 29148629, 265020, 265021, 265022, 65274620, 18108370, 18108374, 264556, 18108385 |
| 1838 | 94324369 (3675, 3676) | Novel Protein sim. GBank gi 1362599 pir J556154 - Abi substrate ena (enabled) - fruit fly (Drosophila melanogaster) | Contains protein domain (PF00568) - WH1 domain | | 29331822, 265017, 264760, 265020, 83373044 |
| 1839 | 87456508 (3677, 3678) | Novel Protein sim. GBank gi 2117310 emb CAB09116.1 - (Z95620) hypothetical protein [Schizosaccharomyces pombe] | | UNCLASSIFIED | 60433438, 264601, 21906765, 21906766, 265021, 33657109, 264556 |
| 1840 | 87391708 (3679, 3680) | Novel Protein sim. GBank gi 127560 sp P23249 MV10_MOUSE - PROTEIN MOV-10 | | UNCLASSIFIED | 264693 |
| 1841 | 85818445 (3681, 3682) | Novel Protein sim. GBank gi 4572464 gb AAD23834.1 AF123653 FEZ1 [Homo sapiens] | | | 56182575, 29331824, 29331826, 60433356, 264764, 264288, 33657023, 263987, 18108370, 18108374, 264631, 264555, 264556, 264639 |

| | | | | | |
|------|-----------------------|---|---|-------------------|--|
| 1842 | 90992645 (3683, 3694) | Novel Protein sim. GBank gi1326268 (U56728) - C54H2.1 gene product [Caenorhabditis elegans] | | UNCLASSIFIED | 65274572, 29331822, 29331824, 29331825, 66714117, 29331826, 29331827, 264907, 264909, 52644045, 56182435, 264510, 265006, 265007, 265009, 264910, 60433356, 264757, 60433438, 55812038, 265017, 18108354, 264686, 264769, 33657023, 264693, 18108364, 33657109, 18108368, 264628, 55810764, 56182323, 18108384, 264563, 264564 |
| 1843 | 95292692 (3685, 3686) | | | UNCLASSIFIED | 264488, 56182435, 264769, 29331826, 29331828, 264511, 265006, 265007, 264910, 264631, 264509, 264690, 264636, 264564, 264691, 60432229, 60432049, 264259, 264629, 33657023, 264486, 264909, 264567, 264595, 264766 |
| 1844 | 87444764 (3687, 3688) | Novel Protein sim. GBank gi2496887[sp]Q09232[YO22 CAEEL - HYPOTHETICAL 32.0 KD PROTEIN C09F5.2 IN CHROMOSOME III | | UNCLASSIFIED | 264908, 265022, 33657023, 87168518, 22279002 |
| 1845 | 95096673 (3689, 3690) | Novel Protein sim. GBank gi1175494[sp]Q09819[YAC5 SCHPO - HYPOTHETICAL 45.0 KD PROTEIN C16C9.05 IN CHROMOSOME I | Contains protein domain (PF00682) - PHD-finger | transcriptfactor | 264259, 29331824, 264907, 264908, 66712502, 264510, 265007, 265008, 55812038, 265018, 21906765, 52644150, 33657109, 264555, 264556, 264557, 56182323, 18108382, 83373044, 18108385, 264564 |
| 1846 | 84287872 (3691, 3692) | Novel Protein sim. GBank gi3881080[emb]CAA21739] - (AL032657) similar to EGF-like domain; cDNA EST yk299a12.3 comes from this gene; cDNA EST EMBL:D35398 comes from this gene; cDNA EST yk331h6.5 comes from this gene; cDNA EST yk299a12.5 comes from this gene; cDNA EST yk467g8.5 c... | Contains protein domain (PF00008) - EGF-like domain | UNCLASSIFIED | 264905, 264908 |
| 1847 | 87821497 (3693, 3694) | Novel Protein sim. GBank gi5059323[gb]AAD38667.1[AF151522] hairy and enhancer of split related-1 [Homo sapiens] | Contains protein domain (PF00010) - Helix-loop-helix DNA-binding domain | transcriptfactor | 22278997, 264259, 29331824, 264909, 18108351, 263974, 22279002 |
| 1848 | 86789360 (3695, 3696) | Novel Protein sim. GBank gi5701854[emb]CAB52191.1] - (AJ245417) G5b protein [Homo sapiens] | | im7 | 29331825, 29331826, 29331827, 265017, 264683, 264288, 264766, 264768, 21906767, 21908768, 264692, 22279002 |
| 1849 | 84287874 (3697, 3698) | Novel Protein sim. GBank gi4503655[ref]NP_001989.1[pFBLN - fibulin 2 precursor | Contains protein domain (PF00008) - EGF-like domain | ATPase_associated | 56182575, 265018 |
| 1850 | 86689650 (3699, 3700) | Novel Protein sim. GBank gi4589582[db]BAA76813.1] - (AB023186) KIAA0969 protein [Homo sapiens] | Contains protein domain (PF00169) - PH domain | strud | 60432049, 264908 |
| 1851 | 95419789 (3701, 3702) | Novel Protein sim. GBank gi220637[db]BAA01471] - (D10827) zinc finger protein [Mus musculus] | Contains protein domain (PF00096) - Zinc finger, C2H2 type | dna_ma_bind | 29331824, 35696052, 264910, 60433438, 264688, 35695917, 265020, 52644150, 65274620, 52644332 |

| | | | | | |
|------|-----------------------|---|--|-------------------|--|
| 1852 | 95413170 (3703, 3704) | Novel Protein sim. GBank gi 5174629 refNP_005090.1 pPIAS - protein inhibitor of activated STAT3 | | UNCLASSIFIED | 56182575, 35696286, 22278996, 22278997, 22278999, 264490, 60432049, 264259, 29331822, 29331824, 29331825, 29331826, 60432289, 29331827, 35696052, 52844045, 265007, 264910, 60433229, 60433358, 60433438, 55812038, 65274444, 265018, 265019, 18108351, 284448, 264686, 264687, 21906765, 21906767, 21906769, 265021, 265022, 52844150, 264693, 33657109, 18108370, 18108374, 55811578, 35695855, 56182323, 60432113, 22279002, 264563 264687, 264768, 52644507, 264769, 21906765, 21906767, 21906768, 22278995, 56994075, 22278999, 52844150, 264259, 264692, 29331822, 29331824, 52845129, 29331827, 33656970, 33657349, 35695763, 264508, 264906, 264628, 264907, 264629, 264909, 35696423, 35695855, 264510, 265006, 264511, 264512, 264630, 265009, 264631, 264910, 264634, 264635, 264637, 264593, 264638, 264639, 33657402, 18108385, 52646317, 52644296, 87168518, 87168559, 264602, 265017, 22279000, 265018, 264760, 264762, 264682, 264448, 264764, 264684, 264567, 264288, 264369, 264766 |
| 1853 | 81222267 (3705, 3706) | Novel Protein sim. GBank gi 854065 emb CAA58337 - (X83413) U88 [Human herpesvirus 6] | | UNCLASSIFIED | 264592 |
| 1854 | 86038152 (3707, 3708) | Novel Protein sim. GBank gi 2072864 (U93569) - putative p150 [Homo sapiens] | | nuclease | |
| 1855 | 81221459 (3709, 3710) | Novel Protein sim. GBank gi 4539520 emb CAB39994.1 - (AL035424) dA22D12.1 (novel protein similar to Drosophila Kelch (Ring Canal protein, KEL) and a heterogeneous set of other types of proteins) [Homo sapiens] | Contains protein domain (PF01344) - Kelch motif | nucl_rept | 18108392, 52646365, 65274572, 56182575, 22278994, 56994075, 22278998, 35696286, 22278999, 264259, 60432049, 264908, 264510, 265007, 265008, 265009, 264595, 21906754, 87168474, 265011, 87168559, 264681, 264288, 264768, 18108359, 21906764, 21906768, 29148627, 265020, 265021, 265022, 52644150, 33657023, 33657109, 18108372, 18108374, 18108376, 35696423, 264631, 264636, 18108381, 264482 |
| 1856 | 94231871 (3711, 3712) | Novel Protein sim. GBank gi 3954978 emb CAA06945 - (AJ008278) acetylglucosaminyltransferase-like protein [Mus musculus] | | UNCLASSIFIED | 56994075, 264259, 29331828, 264511, 264910, 264758, 264693, 264637, 18108381, 83373044 |
| 1857 | 84324455 (3713, 3714) | Novel Protein sim. GBank gi 4322670 gb AAD16120 - (AF094508) dentin phosphoryn [Homo sapiens] | | ATPase_associated | 22278999, 264259, 264906, 60170831, 264448, 264686, 265020, 265022, 33657109, 60170394, 83373044 |

| | | | | | |
|------|-----------------------|--|---|---|--|
| 1858 | 87628311 (3715, 3716) | Novel Protein sim. GBank gij4981903 gb AAD36415.1 AE00178 - (AE001788) ribosomal protein S15 [Thermotoga maritima] | Contains protein domain (PF00312) - Ribosomal protein S15 | 264757 | |
| 1859 | 84407464 (3717, 3718) | Novel Protein sim. GBank gij4240317 bj BAA74937.1 - (AB020721) KIAA0914 protein [Homo sapiens] | | 22278996, 29331824, 265007, 33109954, 265019, 264369, 21906768, 29148784, 27486261, 52644332, 22279002 265019 | |
| 1860 | 17929308 (3719, 3720) | Novel Protein sim. GBank gij4009522 (AF099731) - connexin 31.1 [Homo sapiens] | | | |
| 1861 | 88086370 (3721, 3722) | Novel Protein sim. GBank gij2143637 pir B4505 - calcium- dependent actin-binding protein - rat | Contains protein domain (PF00285) - Citrate synthase | 264887, 264259, 29331822, 29331824, 29331825, 265007, 265009, 264591, 33109954, 265010, 265019, 264369, 264288, 264688, 264691, 264893, 27486264, 18108370, 18108374, 263977, 55811576, 56182323, 264639, 22279000, 22279002, 264482 | |
| 1862 | 87372823 (3723, 3724) | Novel Protein sim. GBank gij125493 sp P07313 KMLC_RABIT - MYOSIN LIGHT CHAIN KINASE, SKELETAL MUSCLE (MLCK) | Contains protein domain (PF00069) - Eukaryotic protein kinase domain | 35696286, 264259, 87168474, 264369, 21908766, 264558, 264583 | |
| 1863 | 85775037 (3725, 3726) | Novel Protein sim. GBank gij3820909 emb CAA09299 - (AJ010642) Dof protein [Drosophila melanogaster] | UNCLASSIFIED | 264601, 264766, 29148627, 29148629, 264692, 264629, 264835 | |
| 1864 | 85547832 (3727, 3728) | Novel Protein sim. GBank gij4322263 gb AAD15985 - (AF077738) metalloproteinase CPX-1 [Mus musculus] | Contains protein domain (PF00754) - F5/8 type C domain | 22278999, 264259, 264907, 265018, 18108370, 264634, 264635, 264555, 264556, 264638, 18108387 | |
| 1865 | 87740827 (3729, 3730) | Novel Protein sim. GBank gij2495727 sp Q93073 Y256_HUMAN - HYPOTHETICAL PROTEIN KIAA0256 | | 22278999, 264490, 29331822, 66714117, 66712502, 265006, 265007, 265008, 265009, 264591, 60433438, 265010, 265019, 264760, 264448, 264768, 29148627, 29148628, 265020, 265022, 18108385, 60432113 | |
| 1866 | 87268816 (3731, 3732) | Novel Protein sim. GBank gij5282617 emb CAB45748.1 - (AL080157) hypothetical protein [Homo sapiens] | kinase | 18108374, 264769, 18108377, 21906765, 21908766, 35696423, 56182575, 21906769, 29148629, 35696286, 35695917, 265021, 264510, 264511, 264512, 264534, 264535, 60170831, 52644150, 264555, 264691, 264259, 264556, 264692, 264557, 33657023, 60433356, 29331822, 264559, 264595, 29331824, 18108385, 21906754, 33657182, 29331827, 35696052, 33656970, 87168518, 265017, 60431602, 22279000, 264508, 264509, 18108351, 264907, 264682, 264567, 18108372, 264765, 264486 | |
| 1867 | 84579159 (3733, 3734) | Novel Protein sim. GBank gij3859830 (AF078096) - forkhead/winged helix-like transcription factor 7 [Homo sapiens] | UNCLASSIFIED | 264094 | |

| | | | | | |
|------|-----------------------|---|--|--------------|--|
| 1868 | 87357459 (3735, 3738) | Novel Protein sim. GBank gi 3881525 emb CAA93884 - (Z70038) cDNA EST EMBL:D32579 comes from this gene; cDNA EST EMBL:D35254 comes from this gene; cDNA EST yk224b3.5 comes from this gene; cDNA EST yk357110.5 comes from this gene [Caenorhabditis elegans] | | nuclease | 264489, 22278997, 22278999, 29331825, 29331826, 265008, 265009, 33657402, 87168474, 18108351, 21906765, 21906768, 21906769, 285020, 265021, 60170615, 27486284, 264628, 18108374, 264631, 18108385, 87168518, 22279000, 22279002, 264566, 264567 |
| 1869 | 86977292 (3737, 3738) | Novel Protein sim. GBank gi 4826772 ref NP_004961.1 pIGFA - insulin-like growth factor binding protein, acid labile subunit | Contains protein domain (PF00560) - Leucine Rich Repeat | glycoprotein | 264508, 264509, 264906, 264908, 264909, 264910, 264591, 264600, 18108351, 264663, 264766, 264769, 35695855, 264634, 264558, 264639, 18108385, 264583, 264486 |
| 1870 | 95349488 (3739, 3740) | Novel Protein sim. GBank gi 18698959 emb CAB06722 - (Z86099) very large tegument protein [human herpesvirus 2] | | UNCLASSIFIED | 29331824, 60424269, 265007, 265008, 21906754, 285017, 265018, 265019, 264288, 264766, 264686, 264688, 21906768, 21906769, 35695917, 60170615, 264692, 18108368, 35695783, 35698423, 65274791, 264638, 264639, 56526486 |
| 1871 | 80234464 (3741, 3742) | | | UNCLASSIFIED | 264509, 264905, 264595, 264768, 264635, 264636, 264563, 264486 |
| 1872 | 80235355 (3743, 3744) | Novel Protein sim. GBank gi 2460316 (AF022147) - uterus-ovary specific putative transmembrane protein [Rattus norvegicus] | | protease | 264510, 264594, 264565 |
| 1873 | 80213890 (3745, 3746) | | | | 264509, 264512, 265009, 265011, 18108351, 264687, 264691, 18108370, 18108374, 264635 |
| 1874 | 95351136 (3747, 3748) | | Contains protein domain (PF00293) - Bacterial mutT protein | UNCLASSIFIED | 264488, 35695917, 264259, 264905, 264907, 264908, 264909, 263978, 264511, 264635, 264636, 264637, 264638, 33657402, 264558, 18108385, 264600, 264604, 264764, 264567, 264766 |
| 1875 | 87330516 (3749, 3750) | Novel Protein sim. GBank gi 4589520 db BAA76782.1 - (AB023155) KIAA0938 protein [Homo sapiens] | | UNCLASSIFIED | 35696286, 264828, 264592, 264557, 264558 |
| 1876 | 87112950 (3751, 3752) | Novel Protein sim. GBank gi 263810 bbs 122920 - collagen alpha chain [Riftia pachyptila=tube worms, Peptide, 1027 aa] | | UNCLASSIFIED | 264259, 29331822, 60432289, 264908, 264909, 264604, 264764, 264288, 264769, 18108376, 264556, 264558, 264559, 18108385 |
| 1877 | 87315208 (3753, 3754) | Novel Protein sim. GBank gi 3983356 gb AAC83924.1 - (AF102545) riboflavin binding protein precursor [Scaphiopus couchii] | | UNCLASSIFIED | 264767, 264686, 264768, 263978, 264693, 264639, 285010, 264563, 264905, 264906, 264907 |

| | | | | | |
|------|-----------------------|---|--|--------------|--|
| 1878 | 95351056 (3755, 3756) | Novel Protein sim. GBank gij4510345jgb AAD21434.1 - (ACO05921) unknown protein [Arabidopsis thaliana] | Contains protein domain (PF01428) - AN1-like Zinc finger | ubiquitin | 264569, 264488, 35696266, 56994075, 264259, 29331822, 29331824, 29331825, 35696052, 29331828, 29146498, 264508, 264509, 264905, 264908, 264907, 264908, 264909, 264510, 264511, 264512, 265009, 264910, 33857402, 264594, 264758, 55812038, 265011, 264602, 264760, 18108351, 264762, 264681, 264682, 264764, 264389, 264288, 264766, 264687, 264768, 264769, 21906766, 35695917, 265021, 60170615, 33657023, 264692, 264693, 33657109, 27486265, 264628, 18108370, 264629, 18108374, 35686423, 264634, 264635, 264555, 264638, 264639, 83373044, 18108385, 56526486, 87168518, 264563, 264564, 264568, 264488, 264567, 264905, 264907, 264908, 265007, 264565, 264566 |
| 1879 | 95310883 (3757, 3758) | Novel Protein sim. GBank gij4929643jgb AAD34082.1 AF15184 - (AF151845) CGI-87 protein [Homo sapiens] | | UNCLASSIFIED | |
| 1880 | 91012978 (3759, 3760) | Novel Protein sim. GBank gij1550785jemb CAA69283 - (Y08026) Immune associated protein 38 [Mus musculus] | | UNCLASSIFIED | 264766, 264691, 264692, 83373044 |
| 1881 | 80214949 (3761, 3762) | Novel Protein sim. GBank gij93144jpir JB40505 - hypothetical protein - suid herpesvirus 1 (strain Indiana-Funkhouser or Becker) | | UNCLASSIFIED | 264509, 264905, 264908, 264909, 264910, 264762, 264687, 33657023, 264632 |
| 1882 | 86582450 (3763, 3764) | Novel Protein sim. GBank gij2384956 (AF022985) - No definition line found [Caenorhabditis elegans] | | | 264908, 21806766, 18108370, 263974, 87168518 |
| 1883 | 94216817 (3765, 3766) | Novel Protein sim. GBank gij1351218jpi P47226 TES2_MOUSE - TESTIN 2 (TES2) [CONTAINS: TESTIN 1 (TES1)] | Contains protein domain (PF00412) - LIM domain containing proteins | homeobox | 264908, 264910, 87168559, 21906766, 264636 |

| | | | | | |
|------|-----------------------|--|--|-------------------|---|
| 1884 | 95310885 (3767, 3768) | Novel Protein sim. GBank gi 4929643 gb AAD34082.1 AF15184 - (AF151845) CGI-87 protein [Homo sapiens] | | UNCLASSIFIED | 264488, 18108394, 56181686, 22278998, 60432049, 264259, 29331822, 29331824, 29331825, 66714117, 29331827, 35696052, 264508, 264509, 264905, 264906, 264907, 264908, 264909, 56182435, 264510, 264511, 265007, 264512, 265009, 264910, 264592, 264593, 33657402, 55812038, 264758, 55811386, 265010, 265011, 87168559, 264600, 265017, 264604, 265019, 264605, 264760, 5581150, 264761, 264882, 264763, 264683, 264764, 264288, 264369, 264766, 264686, 264768, 264769, 29148784, 35695917, 264690, 264691, 33657023, 264692, 264693, 33657109, 18108370, 264628, 60431528, 264629, 263973, 18108374, 55810764, 55811576, 35696423, 35695855, 264631, 264634, 60431850, 264637, 264638, 56182323, 264639, 18108382, 83373044, 18108385, 60432113, 22279002, 264563, 264564, 264565, 264566, 264486, 264567 |
| 1885 | 87644280 (3769, 3770) | Novel Protein sim. GBank gi 2507155 sp P37370 VRP1_YEAST - VERPROLIN | | UNCLASSIFIED | 56182575, 264259, 264905, 264908, 265008, 264596, 264766, 265020, 264628, 60431528, 264634, 56526486, 264080, 264563 |
| 1886 | 86674062 (3771, 3772) | Novel Protein sim. GBank gi 2854158 gb AAC02577.1 - (AF045641) No definition line found [Caenorhabditis elegans] | | | 22278998, 22278999, 60432049, 264910, 265018, 264766, 21906768, 29148629, 264690, 264693, 264628, 264555, 264488 |
| 1887 | 94139139 (3773, 3774) | Novel Protein sim. GBank gi 5174421 ref NP_006023.1 pCPNE - copine VI (neuronal) | Contains protein domain (PF00168) - C2 domain | ATPase-associated | 29331822, 29331824, 29331825, 29331826, 29331827, 264906, 265007, 264681, 264768, 29148627, 264693, 18108364, 35696423, 65274791, 35695855, 264632, 56182323, 264639, 264563 |
| 1888 | 87822804 (3775, 3776) | Novel Protein sim. GBank gi 3319931 emb CAB10841 - (Z98046) dJ1408.2 (Melanoma-Associated Antigen MAGE LIKE) [Homo sapiens] | Contains protein domain (PF01454) - MAGE family | | 263978 |
| 1889 | 91255783 (3777, 3778) | Novel Protein sim. GBank gi 1083308 pir J A56559 - enhancer-trap-locus-1 protein - mouse (fragment) | Contains protein domain (PF00176) - SNF2 and others N-terminal domain | helicase | 264906, 55812038, 264758, 265010, 265018, 265019, 18108351, 264288, 21906765, 21906768, 21906769, 265021, 33657023, 33657109, 56182323, 83373044, 18108385, 22279000, 22279002 |
| 1890 | 87626705 (3779, 3780) | Novel Protein sim. GBank gi 4240185 db BAA74876.1 - (AB020680) KIAA0853 protein [Homo sapiens] | | UNCLASSIFIED | 18108398, 29147620, 264907, 265009, 264600, 265018, 18108351, 264288, 264689, 21906765, 21906768, 21906769, 264691, 264692, 264693, 264628, 18108370, 264638, 264558, 264404 |

| | | | | | |
|------|-----------------------|--|---|--------------|--|
| 1891 | 87013895 (3781, 3782) | | | UNCLASSIFIED | 264686, 264768, 264687, 264692, 264693, 29331822, 29331824, 264508, 264905, 264506, 18108370, 264628, 264907, 264908, 264909, 18108379, 265007, 265008, 264910, 264632, 264591, 264639, 264596, 18108384, 265010, 265011, 264601, 264605, 264563, 264369 |
| 1892 | 87642825 (3783, 3784) | Novel Protein sim. GBank gi 5689535 bj BAA83051.1 - (AB028022) KIAA1099 protein [Homo sapiens] | Contains protein domain (PF01412) - Putative GTP-ase activating protein [for Art] | struct | 22278995, 264509, 87168559, 18108351, 264448, 264682, 265020, 264693, 18108374, 22279000 |
| 1893 | 88533826 (3785, 3786) | | | laminin | 264569, 65274572, 22278997, 22278999, 264259, 29331822, 29331824, 68714117, 29331828, 264906, 265008, 265009, 264592, 265018, 264681, 264448, 264683, 18108354, 264369, 264684, 264685, 264768, 264687, 264689, 21906768, 265020, 265022, 60170615, 52644150, 264690, 264691, 264692, 33657023, 264693, 33657109, 264628, 18108374, 35695855, 264630, 264632, 264634, 264557, 264558, 60170394, 18108381, 18108385, 22279000 |
| 1894 | 86989120 (3787, 3788) | | | | 264508, 264905, 264906, 264907, 264594, 264684, 264690, 264692, 264630, 264635, 264536, 264639, 264583 |
| 1895 | 87631891 (3789, 3790) | Novel Protein sim. GBank gi 5262574 emb CAB45729.1 - (AL080133) hypothetical protein [Homo sapiens] | Contains protein domain (PF00435) - Spectrin repeat | | 56182575, 264258, 60432289, 29331826, 264107, 264905, 264908, 264910, 60170831, 264758, 265010, 265018, 264448, 264288, 264768, 33657109, 264628, 55810764, 18108379, 264634, 56182323, 56526486 |
| 1896 | 85673555 (3791, 3792) | | Contains protein domain (PF00627) - UBA domain | UNCLASSIFIED | 264907, 265008, 264682, 264686, 21906768, 264629, 264631, 264634, 264555 |
| 1897 | 80585589 (3793, 3794) | Novel Protein sim. GBank gi 728836 sp P39193 ALU6_HUMAN - III ALU SUBFAMILY SP WARNING ENTRY III | | cadherin | 264259 |
| 1898 | 87617637 (3795, 3796) | Novel Protein sim. GBank gi 127560 sp P23249 MV10_MOUSE - PROTEIN MOV-10 | | helicase | 22278996, 22278998, 22278999, 29331824, 29331825, 60432289, 29331827, 35696052, 29331828, 265008, 265019, 264681, 264682, 264448, 264369, 52644229, 21906765, 21906766, 21906768, 21906769, 60170615, 55810764, 22279000 |
| 1899 | 86673097 (3797, 3798) | Novel Protein sim. GBank gi 2809819 (AF031548) - erythrocyte membrane glycoprotein Rh50 [Homo sapiens] | Contains protein domain (PF00909) - Ammonium Transporter Family | glycoprotein | 264259, 264508, 264909, 60432229, 264769, 21906765, 21906769 |
| 1900 | 87641858 (3799, 3800) | Novel Protein sim. GBank gi 4102881 (AF017250) - vitellogenin precursor [Oreochromis aureus] | | UNCLASSIFIED | 264683 |

| | | | | | |
|------|-----------------------|--|--|--------------|---|
| 1801 | 95196847 (3801, 3802) | Novel Protein sim. GBank gi 585959 sp P38378 S61A_RAT PROTEIN TRANSPORT PROTEIN SEC61 ALPHA SUBUNIT | Contains protein domain (PF00242) - DNA polymerase (viral) N-terminal domain | transport | 264488, 52644507, 52645156, 18108396, 52646365, 52646842, 18108397, 56182575, 22278894, 22278895, 56994075, 35696286, 22278897, 22278898, 264490, 60432049, 264259, 29331822, 52645080, 29331824, 29331825, 66714117, 29331826, 60432289, 29331827, 29331828, 35696052, 33656970, 29146498, 264906, 264907, 29331830, 264908, 52644045, 264909, 264112, 265006, 264512, 265008, 264910, 265009, 60170831, 60432229, 60433356, 33657402, 60433438, 55812038, 264758, 33109954, 21908754, 33657084, 52644296, 87168474, 265010, 265011, 87168559, 265017, 265018, 265019, 18108351, 264448, 264288, 264888, 52644229, 21908765, 21908766, 21906767, 21906769, 55811957, 35695917, 265020, 265021, 52644150, 18108362, 33657023, 264693, 263987, 33657109, 33657182, 27486264, 33657349, 35695763, 18108370, 18108376, 55811576, 35696423, 35695855, 60431850, 264636, 263981, 52644332, 60170394, 83373044, 18108385, 87168518, 60432113, 264584 264107, 263978 |
| 1802 | 80202013 (3803, 3804) | Novel Protein sim. GBank gi 4426613 gb AAD20451 - (AF098798) SLM-1 [Mus musculus] | | dna_rna_bind | |
| 1803 | 87776554 (3805, 3806) | Novel Protein sim. GBank gi 3747107 (AF095741) - unknown [Rattus norvegicus] | | UNCLASSIFIED | 264759, 29331825, 29331827, 264508, 264907, 265008, 60170831, 60433356, 60433438, 264759, 21908754, 264448, 264288, 265021, 265022, 33657023, 264693, 55811576, 264555, 264556, 22279000 |
| 1804 | 80434213 (3807, 3808) | Novel Protein sim. GBank gi 1352911 sp P47147 YJ80_YEAST - HYPOTHETICAL 80.2 KD PROTEIN IN CPA2-NNF1 INTERGENIC REGION | | struct | 264509, 264905, 264906, 264907, 264908, 265007, 264910, 264686, 264768, 264687, 264769, 264693, 264628, 18108374, 264634, 264636, 264637, 264585 |
| 1805 | 85351140 (3809, 3810) | Novel Protein sim. GBank gi 3043714 dbj BAA25521 - (AB011167) KIAA0595 protein [Homo sapiens] | Contains protein domain (PF00293) - Bacterial mutT protein | | 264488, 264768, 264769, 264689, 29148629, 35695917, 35696286, 264259, 264692, 18108362, 33657023, 29331824, 33657109, 29146499, 264508, 264509, 264905, 264906, 264907, 66712502, 264908, 264909, 35696423, 35695855, 264510, 264511, 264512, 264910, 264634, 264635, 264637, 264638, 33657402, 264758, 85658542, 264802, 264760, 264761, 264482, 264563, 264762, 264483, 264764, 264566, 264288, 264766 |
| 1806 | 12763822 (3811, 3812) | | | UNCLASSIFIED | 264637 |

| | | | | |
|------|-----------------------|--|--|--|
| 1807 | 95351144 (3813, 3814) | Novel Protein sim. GBank gi 4929595 gb AAD34053.1 AF151811 - (AF151816) CGI-58 protein [Homo sapiens] | Contains protein domain (PF00561) - hydrolase alpha/beta hydrolase fold | 65274572, 22278996, 35696286, 22278998, 22278999, 264259, 60432049, 29331822, 29331825, 60424269, 29331828, 35696052, 264593, 60433356, 21906754, 55811386, 85658542, 87168559, 265018, 264681, 264682, 264684, 264288, 21906765, 21906766, 21906768, 265020, 265022, 264890, 52644150, 264692, 33657023, 264693, 33657109, 35695855, 264636, 264638, 60432113 |
| 1808 | 95313641 (3815, 3816) | Novel Protein sim. GBank gi 3986770 (AF109906) - NG22 [Mus musculus] | UNCLASSIFIED | 264488, 65274572, 56182575, 56181686, 22278995, 22278996, 22278997, 22278998, 22278999, 60432049, 264259, 29331822, 60432289, 29331826, 29331827, 29331828, 29146499, 264905, 264906, 264907, 264908, 86712502, 264909, 56182435, 264510, 264511, 265008, 264910, 60170831, 264592, 60433356, 33657402, 264594, 60433438, 264596, 55812038, 33109954, 52646317, 265011, 265017, 264604, 265018, 265019, 264605, 55811150, 264681, 264448, 264288, 264688, 264688, 264769, 21906765, 21906766, 21906767, 21906768, 21906769, 55811957, 29148629, 35695917, 265020, 265022, 264691, 264692, 18108364, 65274620, 33657109, 33657349, 35695763, 18108374, 263978, 55810764, 55811576, 35696423, 65274791, 264631, 264632, 264556, 264557, 60170394, 56182323, 83373044, 18108385, 60432113, 22279000, 22279002, 264566, 264486 |
| 1809 | 85514505 (3817, 3818) | Novel Protein sim. GBank gi 2224653 db BAA20813 - (AB002354) KIAA0356 [Homo sapiens] | UNCLASSIFIED | 264259, 264508, 264905, 264906, 264907, 264908, 264511, 264910, 264593, 264758, 264764, 264766, 18108370, 264634, 264637, 264486 |
| 1910 | 94216821 (3819, 3820) | Novel Protein sim. GBank gi 1351218 sp P47226 TES2_MOUSE - TESTIN 2 (TES2) [CONTAINS: TESTIN 1 (TES1)] | Contains protein domain (PF00412) - LIM domain containing proteins | 35696286, 22278996, 22278999, 35696052, 264509, 264905, 264906, 264907, 264908, 264909, 264511, 265007, 264512, 264910, 264758, 265011, 264601, 264602, 264604, 264605, 264761, 264764, 264288, 264768, 264768, 264687, 264769, 35695917, 265021, 52644150, 264692, 264628, 18108370, 264629, 18108372, 18108374, 35695855, 264631, 264634, 264635, 264636, 264637, 264638, 18108385, 264565, 264566, 264486 |

| | | | | | |
|------|-----------------------|--|--|--------------|---|
| 1811 | 91725345 (3821, 3822) | Novel Protein sim. GBank gi 4809339 gb AAD30184.1 AC006530 - (AC006530) hypothetical protein [Homo sapiens] | Contains protein domain (PF01119) - DNA mismatch repair protein | nuclease | 18108394, 56182575, 56182181, 29331826, 29331827, 33656970, 264906, 265007, 264591, 55812038, 87168559, 264448, 264389, 21906765, 21906768, 265022, 264691, 264693, 18108385, 55811576, 264556, 18108385, 18108388 |
| 1812 | 95413519 (3823, 3824) | Novel Protein sim. GBank gi 5689439 dbj BAA83003.1 - (AB028974) KIAA1051 protein [Homo sapiens] | Contains protein domain (PF00098) - Zinc finger, CCHC class | UNCLASSIFIED | 18108397, 56182575, 56181886, 22278894, 22278995, 56994075, 22278996, 22278999, 264259, 29331822, 29331824, 56182181, 29331825, 66714117, 35696052, 264605, 264906, 264907, 264908, 52844045, 56182435, 265007, 265008, 264910, 265009, 264591, 264596, 65274444, 55811386, 87168474, 265011, 87168559, 265018, 265019, 264760, 18108351, 264681, 264369, 264684, 264288, 264686, 264768, 21906765, 21906766, 21906767, 21906768, 21906769, 55811957, 265020, 265021, 265022, 60170615, 264692, 33657023, 264693, 18108376, 55811576, 35696423, 65274791, 264637, 56182323, 83373044, 56526486, 22279002, 264563, 264566 |
| 1813 | 95305546 (3825, 3826) | Novel Protein sim. GBank gi 5032245 ref NP_005665.1 pZNF2 - zinc finger protein (C2H2) homologous to mouse MOK-2 | | | 56182575, 22278994, 22278995, 56994075, 22278996, 22278998, 22278999, 29331826, 29331827, 265006, 55812038, 265010, 265017, 265018, 265019, 264681, 18108351, 264683, 264764, 264369, 264288, 264685, 264686, 264769, 21906765, 21906766, 21906768, 21906769, 55811957, 265020, 265022, 264691, 55811576, 264634, 264635, 264638, 56182323, 83373044, 18108385 |
| 1814 | 83423982 (3827, 3828) | Novel Protein sim. GBank gi 4589604 dbj BAA76824.1 - (AB023197) KIAA0980 protein [Homo sapiens] | Contains protein domain (PF00036) - EF hand | struct | 56182575, 29331824, 35696052, 264906, 264908, 264828, 264909, 264592, 264758, 87168559, 18108351, 18108354, 264684, 264686, 33657023, 264693, 264628, 264631, 264632, 264634, 264635, 264639 |
| 1815 | 95340459 (3829, 3830) | Novel Protein sim. GBank gi 5689415 dbj BAA82991.1 - (AB028962) KIAA1039 protein [Homo sapiens] | | UNCLASSIFIED | 264259, 29331824, 29331826, 29331827, 264508, 264909, 265009, 265017, 265019, 264768, 264769, 264689, 264628, 264635, 264637, 264639, 83373044, 264585 |
| 1816 | 79640761 (3831, 3832) | | | | 264693, 264639 |

| | | | | | |
|------|-----------------------|--|--------------|--------|--|
| 1917 | 87821680 (3833, 3834) | Novel Protein sim. GBank gi15669391 dbj BAA82979.1 - (AB028950) KIAA1027 protein [Homo sapiens] | | struct | 284769, 284688, 21806765, 21806768, 22278996, 264259, 264691, 264693, 29331824, 29331825, 29331826, 29331828, 264905, 264908, 264628, 264907, 264908, 264909, 264510, 264630, 264910, 264634, 264635, 264638, 264637, 264638, 263981, 264639, 264758, 18108385, 21806754, 265011, 264604, 264563, 18108351, 264762, 264763, 264568, 264764, 264766 |
| 1918 | 95302785 (3835, 3836) | Novel Protein sim. GBank gi15281517 gb AAD41524.1 AF15483 - (AF154831) PV-1 [Rattus norvegicus] | | struct | 284488, 18108392, 18108357, 21806765, 21906767, 21906768, 56182575, 21906769, 22278994, 35696286, 35695917, 22278996, 22278997, 265021, 265022, 264534, 264690, 264691, 264692, 33857023, 264693, 29331824, 29331825, 33857109, 29331828, 52845129, 35696052, 29331828, 27486262, 27486264, 35695763, 284508, 264905, 284509, 264908, 264628, 264907, 18108370, 264908, 264629, 264909, 18108372, 18108374, 263978, 35696423, 35695855, 264510, 264511, 265006, 265007, 264512, 265008, 264631, 265009, 264910, 264634, 264635, 264555, 264636, 264556, 264637, 264557, 264593, 264638, 264594, 60170394, 264595, 264559, 264596, 83373044, 264758, 52846317, 18108385, 52844286, 56528486, 87168518, 265010, 265011, 87168559, 264600, 264601, 264602, 265017, 264603, 264604, 265018, 264605, 264760, 264761, 264482, 264564, 18108351, 264762, 264682, 264565, 264448, 264764, 264566, 264486, 264567, 264389, 264288, 264766, 264487, 264685 |
| 1919 | 94143847 (3837, 3838) | Novel Protein sim. GBank gi13878584 emb CAB01237 - (Z77667) cDNA EST EMBL:C08125 comes from this gene; cDNA EST EMBL:C08753 comes from this gene [Caenorhabditis elegans] | oxidase | | 22278997, 29331822, 265007, 60170831, 60432229, 60433438, 264448, 264682, 264288, 55811957, 33857023, 33857109, 65274791, 56182323, 22279002, 264510, 264511, 264512, 264566 |
| 1920 | 91229853 (3839, 3840) | Novel Protein sim. GBank gi11809231 (AC000115) - coded for by human cDNAs R76043 (NID:g850725), R65857 (NID:g838495) and H12868 (NID:g877688) [Homo sapiens] | UNCLASSIFIED | | |
| 1921 | 79555226 (3841, 3842) | Novel Protein sim. GBank gi14580997 gb AAD24571.1 AF12108 - (AF121081) cAMP inducible 2 protein [Mus musculus] | UNCLASSIFIED | | 264693 |

| | | | | | |
|------|-----------------------|--|---|----------------|--|
| 1922 | 87641863 (3843, 3844) | Novel Protein sim. GBank gij138593spP02845VIT2, CHICK - VITELLOGENIN II PRECURSOR (MAJOR VITELLOGENIN) [CONTAINS: LIPOVITELLIN I (LV); PHOSVITIN (PV); LIPOVITELLIN II (LVII); YGP40] | | UNCLASSIFIED | 264686, 264688, 264490, 18108370, 264909, 18108374, 265008, 264557, 264564, 18108351 |
| 1923 | 94323589 (3845, 3846) | Novel Protein sim. GBank gij119110spP03211 EBN1_EBV - EBNA-1 NUCLEAR PROTEIN | | UNCLASSIFIED | 264488, 264489, 22278995, 264094, 264259, 35696052, 264509, 264905, 264906, 264907, 264908, 66712502, 264908, 264511, 264512, 265007, 264910, 265009, 264591, 264592, 264593, 264594, 264595, 264758, 264596, 264759, 265011, 265017, 265018, 265019, 55811150, 264681, 264762, 264448, 264764, 264288, 264369, 264766, 264767, 264686, 264687, 264768, 264769, 56181562, 264689, 21906766, 264691, 33657023, 264693, 65274620, 33657109, 18108370, 264628, 264629, 35698423, 264630, 264631, 264632, 264634, 264555, 264636, 264637, 264638, 264639, 264558, 83373044, 87168518, 264563, 264564, 264565, 264566, 264567 |
| 1924 | 87338925 (3847, 3848) | Novel Protein sim. GBank gij387655 emb CAA96657 - (Z72511) possible zinc finger protein; cDNA EST EMBL: M89115 comes from this gene; cDNA EST EMBL: D71533 comes from this gene; cDNA EST EMBL: D72314 comes from this gene; cDNA EST EMBL: D75164 comes from this gene; cDNA EST EMBL: C1... | Contains protein domain (PF00097) - Zinc finger, C3HC4 type (RING finger) | | 29331826, 264906, 264908, 264595, 265017, 265018, 265019, 265021, 264691, 264693, 264637, 18108385, 264565 |
| 1925 | 87628338 (3849, 3850) | Novel Protein sim. GBank gij4981903 gb AAD36415.1 AE001788 ribosomal protein S15 [Thermotoga maritima] | Contains protein domain (PF00312) - Ribosomal protein S15 | ribosomal prot | 22278995, 22278996, 22278997, 264259, 29331824, 66714117, 29146499, 264909, 52644045, 265008, 265009, 264758, 265011, 265017, 264605, 264448, 264288, 264692, 33657109, 18108374, 60170394 |
| 1926 | 88084739 (3851, 3852) | Novel Protein sim. GBank gij2246532 (U93872) - ORF 73, contains large complex repeat CR 73 [Kaposi's sarcoma- associated herpesvirus] | | strud | 264905, 264908, 264907, 264910, 264593, 265018, 264760, 264764, 264288, 264692, 264693, 263978, 264631, 264634, 264637, 264563 |
| 1927 | 85654657 (3853, 3854) | Novel Protein sim. GBank gij3043632 dbj BAA25480 - (AB011126) KIAA0554 protein [Homo sapiens] | | | 264689, 264631 |
| 1928 | 87788054 (3855, 3856) | Novel Protein sim. GBank gij166576 dbj BAA13377 - (D87433) KIAA0246 [Homo sapiens] | Contains protein domain (PF00193) - Inf | Inf | 264489, 264259, 265017, 265021, 264692 |
| 1929 | 86997236 (3857, 3858) | Novel Protein sim. GBank gij5001993 gb AAD37247.1 AF13432 - (AF134321) chimeric AFGPtypsinogen-like serine protease precursor [Dissostichus mawsoni] | UNCLASSIFIED | UNCLASSIFIED | 264508, 264591, 33657402, 265017, 264768, 264632, 264556, 264639 |

| | | | | | | |
|------|-----------------------|--|--|---|--------------|--|
| 1830 | 87889128 (3859, 3860) | Novel Protein sim. GBank gi 1709230 sp P52963 NBL4_MOUSE - NBL4 PROTEIN | | | phosphatase | 35696286, 29331828, 264905, 264907, 264908, 264909, 264511, 264910, 264758, 264601, 265017, 265019, 264605, 264760, 264764, 264766, 264688, 264769, 265022, 35698423, 264638, 60432113 |
| 1831 | 87797279 (3861, 3862) | Novel Protein sim. GBank gi 404634 (U01840) - serine/threonine kinase [Mus musculus] | | Contains protein domain (PF00069) - Eukaryotic protein kinase domain | kinase | 264906, 264908, 60432229, 264758, 264764, 264288, 265020, 264692, 264634, 264637, 264684, 264691, 264635 |
| 1832 | 15030972 (3863, 3864) | | | | UNCLASSIFIED | 264595 |
| 1833 | 11813668 (3865, 3866) | Novel Protein sim. GBank gi 4115748 gb BAA36494 - (AB022023) nonmuscle myosin heavy chain B [Bos taurus] | | | struct | 56182575, 56182435, 264510, 264757, 264758, 55812038, 55811388, 265018, 55811150, 21906765, 264691, 264631, 264635, 264637 |
| 1835 | 87752511 (3869, 3870) | | | | UNCLASSIFIED | 264688, 265011, 264511, 264905, 18108351, 264564, 264681, 264259, 18108370, 264566, 264764, 264359, 264595 |
| 1836 | 85414336 (3871, 3872) | Novel Protein sim. GBank gi 4827040 ref NP_005110.1 pTRAP - thyroid hormone receptor-associated protein, 150 kDa subunit | | | | 60432289, 265007, 265010, 265011, 265018, 33657109, 18108374 |
| 1837 | 84847141 (3873, 3874) | Novel Protein sim. GBank gi 543187 pir S37771 - ankryn, erythrocyte - mouse | | Contains protein domain (PF00023) - Ank repeat | kinase | 85658542, 21906767, 35695917, 60170615, 264693, 33657109 |
| 1838 | 87403277 (3875, 3876) | Novel Protein sim. GBank gi 4544431 gb AAD22340.1 AC006955 - (AC006955) hypothetical protein [Arabidopsis thaliana] | | Contains protein domain (PF00888) - Cullin family | collagen | 264488, 28146498, 264905, 264559 |
| 1839 | 91004978 (3877, 3878) | Novel Protein sim. GBank gi 500858 db BAA03210 - (D14168) 50kDa lectin [Bombyx mori] | | | UNCLASSIFIED | 85274572, 22278999, 60432289, 56182435, 60433356, 265017, 21906765, 21906766, 21906768, 55811957, 27486264, 35698423, 60432113, 264584 |
| 1840 | 87348810 (3879, 3880) | Novel Protein sim. GBank gi 1946300 emb CAA73132 - (Y12529) hypothetical protein [Silene latifolia] | | Contains protein domain (PF00560) - Leucine Rich Repeat | struct | 264488, 28331822, 264448, 264683, 264288, 265020, 33657023, 264631 |
| 1841 | 94147177 (3881, 3882) | Novel Protein sim. GBank gi 4208386 (AF060570) - rig-1 protein [Mus musculus] | | | UNCLASSIFIED | 56994075, 22278998, 264259, 29331824, 29331827, 264905, 265008, 33657084, 265017, 265018, 264288, 264687, 21906765, 21906766, 21906767, 265020, 52644150, 27486264, 83373044, 18108387, 60432113, 22279002, 264565 |
| 1842 | 87641870 (3883, 3884) | Novel Protein sim. GBank gi 4927204 gb AAD33049.1 AF133911 - (AF133911) ARL-6 interacting protein-4 [Mus musculus] | | | UNCLASSIFIED | 264488, 18108398, 28331825, 27486261, 264509, 18108370, 18108374, 264482 |
| 1843 | 94325298 (3885, 3886) | Novel Protein sim. GBank gi 3122952 sp O15736 TIPD_DICD1 - TIPD PROTEIN | | Contains protein domain (PF00400) - WD domain, G-beta repeat | kinase | 22278998, 29331822, 29331827, 35696052, 264511, 265009, 264592, 60432229, 265017, 265018, 265019, 264684, 264692, 33657109, 85274791, 264636 |

| | | | | |
|------|-----------------------|--|--------------|--|
| 1844 | 94232958 (3887, 3888) | Novel Protein sim. GBank gi 1799570 dbj BAA13432 - (D87671) TIP120 [Rattus norvegicus] | UNCLASSIFIED | 65274572, 22278994, 22278995, 22278996, 22278998, 22278999, 264259, 52645080, 29331824, 29331826, 29331827, 66712502, 56182435, 264512, 264910, 60170831, 60432229, 60433358, 33657402, 60433438, 264586, 33109954, 21906754, 87168474, 87168559, 265017, 265018, 265019, 18108351, 264369, 264686, 264768, 21906765, 21906768, 21906767, 21906769, 35695917, 265020, 265021, 60170615, 264692, 33657023, 18108370, 18108374, 35696423, 35695855, 264634, 60170394, 264639, 83373044, 18108385, 56526486, 87168518, 60432113, 22279000, 264563, 264488, 22278996, 264510, 264511, 18108351, 264683, 264486, 264567 |
| 1845 | 87641872 (3889, 3890) | Novel Protein sim. GBank gi 4927204 gb AAD33048.1 AF13391 - (AF133911) ARL-6 interacting protein-4 [Mus musculus] | UNCLASSIFIED | 60432289, 29331827, 35696052, 265007, 265008, 60433356, 60433438, 264369, 56181562, 21906767, 52844150, 264693, 27486264, 264637, 87168518, 264563 |
| 1846 | 87443980 (3891, 3892) | Novel Protein sim. GBank gi 2498104 sp Q27969 AD50_BOVIN - ADRENAL MEDULLA 50 KD PROTEIN | UNCLASSIFIED | 22278998, 264905, 264906, 264908, 264909, 264512, 264758, 264762, 264682, 264683, 264764, 264288, 264768, 264688, 21906768, 264693, 18108374, 35695855, 264635, 264637, 264638, 18108385, 22279002 |
| 1847 | 86438862 (3893, 3894) | Novel Protein sim. GBank gi 3914801 sp O54888 RPA2_RAT - DNA-DIRECTED RNA POLYMERASE I 135 KD POLYPEPTIDE (RNA POLYMERASE I SUBUNIT 2) (RPA135) (RNA POLYMERASE I 127 KD SUBUNIT) | mapolymerase | 264909, 60170831, 264591, 264594, 235010, 265011, 264764, 264369, 264689, 264631, 264638 |
| 1848 | 85189174 (3895, 3896) | Novel Protein sim. GBank gi 5420387 emb CAB46679.1 - (AJ243458) proteophosphoglycan [Leishmania major] | struct | 264369 |
| 1849 | 7640129 (3897, 3898) | Novel Protein sim. GBank gi 3876766 emb CAA93466.1 - (Z69637) predicted using GeneFinder; Similarity to E.coli hypothetical protein YCAC (SW:YCAC_ECOLI) [Caenorhabditis elegans] | UNCLASSIFIED | 264488, 264768, 264688, 264689, 264259, 29331822, 33657109, 29331828, 264103, 264509, 18108370, 35695855, 264510, 265008, 265009, 33657402, 18108385, 265018, 264583, 55811150, 18108351, 264369, 264288, 18108354 |
| 1850 | 87786531 (3899, 3900) | Novel Protein sim. GBank gi 2626753 dbj BAA23424 - (AB008782) sulfate transporter [Arabidopsis thaliana] | transport | 56182575, 22278997, 52645080, 29331824, 29331825, 29331827, 55812038, 52846317, 265018, 265019, 264369, 21906765, 21906767, 55811957, 265020, 265021, 33657023, 264693, 35695763, 58182323, 22279002 |
| 1851 | 86988253 (3901, 3902) | Novel Protein sim. GBank gi 4929633 gb AAD34077.1 AF15184 - (AF151840) CGI-82 protein [Homo sapiens] | reductase | 264591, 264593, 264594, 264595, 264555, 264556, 264557, 264558, 264565 |

| | | | | | |
|------|-----------------------|---|--|--------------|---|
| 1953 | 20470371 (3905, 3906) | Novel Protein sim. GBank gi 168715 sp P31721 C1QB_RAT - COMPLEMENT C1Q SUBCOMPONENT, B CHAIN PRECURSOR Novel Protein sim. GBank gi 4240271 dbj BAA74914.1 - (AB020698) KIAA0891 protein [Homo sapiens] | Contains protein domain (PF00386) - C1q domain | complement | 264259, 264558 |
| 1954 | 91226025 (3907, 3908) | | Contains protein domain (PF00443) - Ubiquitin carboxyl-terminal hydrolase family 2 | ubiquitin | 264488, 263994, 18108394, 35696286, 22278998, 29331822, 66714117, 29331828, 29331827, 264508, 264509, 264905, 264906, 264907, 264908, 264909, 264510, 265008, 265007, 265008, 264910, 265009, 60170831, 60432229, 265011, 265017, 264603, 264604, 265019, 18108351, 264685, 264766, 264769, 35895917, 265020, 265021, 264691, 33637023, 264692, 33637109, 264628, 18108374, 35696423, 55811576, 35695855, 264630, 264635, 264636, 264555, 264556, 264638, 264557, 264639, 264558, 264559, 93373044, 18108385, 264563, 264564, 264566, 264488, 264567 |
| 1955 | 95308310 (3909, 3910) | | | UNCLASSIFIED | 56182575, 56181686, 22278994, 22278999, 264259, 29331822, 56182181, 29331824, 29331825, 29331826, 29331827, 35696052, 264508, 29331830, 265008, 265009, 264591, 55812038, 87168474, 265017, 265018, 265019, 264448, 264766, 21906765, 21906766, 21906767, 55811957, 265020, 265021, 52645129, 33637109, 27486264, 33637349, 35695763, 60431528, 18108374, 55811576, 35695855, 264635, 60431850, 264639, 83373044 |
| 1956 | 95092121 (3911, 3912) | Novel Protein sim. GBank gi 1665821 dbj BAA134071 - (D87469) Similar to D.melanogaster cadherin-related tumor suppressor [Homo sapiens] | Contains protein domain (PF00028) - Cadherin domain | cadherin | 264488, 56182575, 22278996, 22278999, 264259, 29331822, 29331824, 29331825, 29331827, 35696052, 29331828, 264508, 264906, 264908, 264909, 264511, 265007, 264910, 264591, 55812038, 265010, 265018, 18108351, 264768, 56181562, 264689, 21906768, 21906769, 265022, 264691, 264628, 18108374, 55810764, 55811576, 35695855, 264631, 264632, 264635, 264637, 264639, 60170394, 56182323, 83373044, 18108385, 22279000, 22279002, 264563, 264564 |

| | | | | | |
|------|-----------------------|---|--|--------------|--|
| 1957 | 94326510 (3913, 3914) | Novel Protein sim. GBank gi4589674[dbj BAA76856.1] - (AB023229) KIAA1012 protein [Homo sapiens] | | UNCLASSIFIED | 52648642, 56182575, 22278997, 22278998, 22278999, 29331824, 66714117, 29331827, 29146498, 264593, 33657402, 33109954, 87169474, 265018, 264448, 264369, 264288, 264766, 21906766, 21906767, 21906768, 21906769, 265020, 265021, 264692, 65274620, 27486264, 33657349, 27486265, 33695855, 22279002, 264482 |
| 1958 | 95313902 (3915, 3916) | Novel Protein sim. GBank gi4240227[dbj BAA74892.1] - (AB020676) KIAA0869 protein [Homo sapiens] | | UNCLASSIFIED | 22278998, 264092, 264084, 264259, 80432048, 29331824, 56182181, 86714117, 264107, 264109, 264909, 264511, 60170831, 60432229, 21906754, 265010, 21906769, 33695917, 265022, 65274620, 263967, 263976, 35698423, 264631, 264632, 264634, 264635, 18108385, 22279000, 22279002, 264593, 265019 |
| 1959 | 85701470 (3917, 3918) | Novel Protein sim. GBank gi2281983[emb CAB10860] - (Z98056) hyppothetical protein [Schizosaccharomyces pombe] | | ubiquitin | |
| 1960 | 80308608 (3919, 3920) | Novel Protein sim. GBank gi2274851[dbj BAA21515] - (D64159) 3-7 gene product [Homo sapiens] | | struct | 264905, 264906, 264907, 264908, 264909, 265006, 265007, 264910, 264595, 265017, 264604, 265018, 18108351, 264764, 264369, 264766, 264768, 21906765, 18108368, 264629, 18108379, 264635, 264636, 264637, 264638, 264486 |
| 1961 | 16292607 (3921, 3922) | | | | 264635 |
| 1962 | 91008385 (3923, 3924) | | | UNCLASSIFIED | 65274572, 264592, 264593, 265019, 264691 |
| 1963 | 80938017 (3925, 3926) | Novel Protein sim. GBank gi3721653[dbj BAA33581] - (AB012933) acyl-CoA synthetase 5 [Rattus norvegicus] | | eph | 65274572, 18108398, 35695286, 29331825, 60432289, 29331827, 264828, 265008, 265009, 60433356, 80433438, 21906754, 265020, 265021, 33657023, 33657109, 27486265, 35695855, 264555 |
| 1964 | 94317605 (3927, 3928) | Novel Protein sim. GBank gi5262638[emb CAB45757.1] - (AL080169) hypothetical protein [Homo sapiens] | | cadherin | 264488, 264092, 264259, 264509, 264905, 264906, 264907, 264908, 264909, 264510, 264511, 265007, 265009, 264910, 264592, 264593, 264594, 264595, 264758, 264600, 264603, 264604, 264805, 264760, 264762, 264448, 264764, 264288, 264885, 264766, 264768, 264769, 21906766, 264691, 264692, 264693, 18108370, 264628, 264629, 18108374, 264630, 264631, 264634, 264636, 264637, 264638, 18108382, 83373044, 18108385, 264483, 264564, 264565, 264566, 264486, 264567 |
| 1965 | 94317445 (3929, 3930) | Novel Protein sim. GBank gi4107017[dbj BAA36294] - (AB001773) PEM-6 [Ciona savignyi] | Contains protein domain (PF01428) - AN1-like Zinc finger | ubiquitin | 264488, 264510, 264760, 264768, 264486 |

| | | | | | |
|------|-----------------------|---|---|--------------|---|
| 1866 | 84192058 (3931, 3932) | Novel Protein sim. GBank gij4929707igbAAD34114.1 AF15187 - (AF15187) CGI-119 protein [Homo sapiens] | Contains protein domain (PF01027) - Uncharacterized protein family | glycoprotein | 22278999, 264092, 264259, 29331826, 29331828, 29146488, 264595, 265011, 264448, 18108354, 264288, 264684, 264766, 264685, 264686, 265022, 264691, 264692, 18108370, 18108377, 264555, 18108381, 18108385, 264486, 264587 |
| 1867 | 87395123 (3933, 3934) | Novel Protein sim. GBank gij2957270 (AF044576) - phospholipase C PLC210 [Caenorhabditis elegans] | Contains protein domain (PF00388) - Phosphatidylinositol-specific phospholipase C, x domain | esterase | 29331824, 265010, 265017, 264288, 21906764, 263981, 56526486 |
| 1868 | 88095641 (3935, 3936) | Novel Protein sim. GBank gij2564953 (AF030001) - unknown [Mus musculus] | Contains protein domain (PF00008) - EGF-like domain | oncogene | 35698286, 264905, 264509, 264906, 264907, 264908, 264909, 264511, 264512, 265006, 264910, 265009, 264594, 264757, 264758, 264604, 264605, 264760, 264762, 264682, 264764, 264685, 264766, 264767, 264689, 264691, 264693, 264628, 264629, 35698423, 35695855, 264631, 264632, 264634, 264635, 264636, 264637, 18108380, 264564, 264565, 264566, 264567 |
| 1869 | 84328529 (3937, 3938) | Novel Protein sim. GBank gij2911274 (U20329) - spidroin 1 [Nephila clavipes] | | UNCLASSIFIED | 22278995, 22278996, 35696052, 264906, 264908, 18108351, 264482 |
| 1870 | 80596049 (3939, 3940) | Novel Protein sim. GBank gij4050087 (AF109907) - S164 [Homo sapiens] | | UNCLASSIFIED | 264908, 264288, 264766, 264636 |
| 1871 | 94843914 (3941, 3942) | Novel Protein sim. GBank gij134206 spjP09593 SANT_PLAFV - S-ANTIGEN PROTEIN PRECURSOR | | collagen | 264488, 264489, 22278998, 264259, 60432049, 66714117, 29331826, 60432289, 29331827, 35696052, 264508, 264905, 264508, 264906, 264907, 264908, 264909, 264510, 264511, 264512, 264910, 264591, 264592, 60432229, 60433356, 264595, 264596, 264600, 264604, 264605, 264760, 18108351, 264448, 264764, 264288, 264766, 264768, 264769, 21906765, 33657023, 264692, 18108370, 264629, 35696423, 65274791, 35695855, 264632, 264635, 264555, 264636, 264637, 264638, 264639, 18108385, 60432113, 22278000, 264563, 264564, 264565, 264566, 264486 |
| 1872 | 87645444 (3943, 3944) | Novel Protein sim. GBank gij4519623 dbj BAA75671.1 - (AB017616) homologous to the yeast YGR163 gene [Mus musculus] | Contains protein domain (PF01462) - Leucine rich repeat N-terminal domain | | 22278999, 264259, 29331822, 56182181, 60432289, 29331827, 52644045, 264909, 265006, 264511, 265008, 52644296, 265018, 265019, 264761, 264689, 21906768, 21906769, 264691, 264693, 33657109, 33657182, 264556, 52644332, 264558, 60432113 |
| 1873 | 86395533 (3945, 3946) | | | UNCLASSIFIED | 28331828, 264692, 35696423, 264631, 264555, 264556, 264557, 264558, 264559 |
| 1874 | 80396629 (3947, 3948) | Novel Protein sim. GBank gij3309543 (AF036382) - MLL [Fugu rubripes] | | UNCLASSIFIED | 264682, 264764, 264563 |

| | | | | | |
|------|-----------------------|---|--|----------------------|--|
| 1975 | 94316479 (3949, 3950) | | | UNCLASSIFIED | 264488, 66714117, 29331826, 29331828, 56182435, 265006, 264757, 55812038, 265010, 265017, 264369, 55811957, 65274781, 35695855, 56182323, 60432113, 264258, 35696052, 265018, 265020, 265021, 33657109, 56526486 |
| 1976 | 95358914 (3951, 3952) | | | UNCLASSIFIED | 264908, 264596, 265021, 264566 |
| 1977 | 94852664 (3953, 3954) | Novel Protein sim. GBank gij2499526ipIQ07782INASU_RAT - SODIUM/SULFATE COTRANSPORTER (NA(+)/SULFATE COTRANSPORTER) | | homeobox | |
| 1978 | 87447645 (3955, 3956) | Novel Protein sim. GBank gij103421ipijA33471 - transcription factor NTF1 - fruit fly (Drosophila melanogaster) (fragment) | | transcription factor | 60170831, 264568 |
| 1979 | 87627709 (3957, 3958) | Novel Protein sim. GBank gij2244815lembjCAB10238.1] - (Z97338) hypothetical protein (Arabidopsis thaliana) | | UNCLASSIFIED | 29331826, 28146498, 264905, 264907, 265007, 265009, 265010, 265018, 264686, 18108359, 21908768, 35695917, 265020, 60170615, 264693, 18108368, 18108370, 264631, 264635, 264556, 264558, 18108384, 22279000, 264565 |
| 1980 | 86577059 (3959, 3960) | Novel Protein sim. GBank gij4759290jefINP_004642.1pUSP1 - Ubiquitin carboxyl-terminal hydrolase, X-linked | | ubiquitin | 264489 |
| 1981 | 87606974 (3961, 3962) | Novel Protein sim. GBank gij4406683jgbjAAD20060] - (AF131849) Unknown (Homo sapiens) | | UNCLASSIFIED | 22278995, 22278997, 22278999, 264259, 29331825, 29331828, 29146498, 29146499, 264107, 264908, 264910, 264595, 21906754, 265010, 265017, 265018, 265019, 264448, 264288, 21908767, 33657023, 27486264, 18108370, 18108374, 18108376, 264630, 264631, 264635, 18108385, 87168518, 22279000, 264482, 264564 |
| 1982 | 90995367 (3963, 3964) | Novel Protein sim. GBank gij5689523dbjBAA83045.1] - (AB029016) KIAA1093 protein (Homo sapiens) | | | 65274572, 29331825, 35696052, 33656970, 264909, 265008, 55811386, 264760, 264686, 264691, 27486264 |
| 1983 | 95098688 (3965, 3966) | Novel Protein sim. GBank gij3417297 (AC002310) - Unknown gene product (Homo sapiens) | Contains protein domain (PF00096) - Zinc finger, C2H2 type | transcription factor | 22278996, 35696286, 264259, 29331822, 20281099, 29331824, 60432289, 29331827, 264509, 264905, 264906, 264907, 66712502, 264908, 52844045, 264909, 264510, 264512, 264910, 265009, 264591, 264592, 60433356, 60433438, 264758, 265010, 264600, 264603, 264604, 264760, 264762, 264763, 264764, 264768, 264687, 264768, 264769, 21908765, 55811957, 35695917, 264690, 264692, 264693, 264628, 264629, 263978, 18108379, 35696423, 35695855, 20281071, 264632, 264634, 264635, 264555, 264636, 264637, 264638, 264639, 83373044, 18108385, 22279000, 22279002 |

| | | | | | |
|------|-----------------------|--|--|--------------|--|
| 1884 | 85760989 (3867, 3968) | Novel Protein sim. GBank gij2896695[emb]CAA17174.1] - (AL021887) fadD14 [Mycobacterium tuberculosis] | | synthase | 264688, 21906766, 55811957, 56994075, 265020, 265021, 22278999, 265022, 264259, 29331822, 33657182, 29148499, 264628, 18108370, 264908, 264629, 55811576, 35695855, 265008, 265007, 264591, 21908754, 33657084, 265010, 265017, 265019, 264288 |
| 1885 | 85636897 (3968, 3970) | Novel Protein sim. GBank gij5712131[gb]AAD47379.1[AF12049] DEM1 protein [Homo sapiens] | | glycoprotein | 264760, 264288, 263978, 55811576, 264637, 56182323, 18108385, 264564 |
| 1886 | 80200507 (3971, 3972) | | | UNCLASSIFIED | 264488, 264629 |
| 1887 | 87011117 (3973, 3974) | Novel Protein sim. GBank gij4868443[gb]AAD31319.1[AF14457] - (AF144573) Mx-interacting protein kinase PKM [Mesocricetus auratus] | Contains protein domain (PF00089) - Eukaryotic protein kinase domain | | 22278999, 29331830, 265007, 265018, 21906768, 33657023, 264692, 264693, 18108377, 264635, 60170394, 22279002 |
| 1888 | 94122108 (3975, 3976) | | | UNCLASSIFIED | 264905, 264906, 264907, 264908, 264909, 264910, 264591, 264593, 264758, 264764, 264686, 264768, 265021, 264692, 264628, 264629, 35695855, 264630, 264635, 264636, 264637, 264638, 264639, 264483 |
| 1889 | 91225225 (3977, 3978) | Novel Protein sim. GBank gij2801701 (AF042379) - spindle pole body protein spc97 homolog GCP2 [Homo sapiens] | | tubulin | 60432049, 60432289, 52644045, 56182435, 264112, 265007, 33657402, 52644229, 21906765, 21906768, 21906769, 55811957, 33657023, 263967, 33657109, 18108370, 22279000, 22279002 |
| 1890 | 85699888 (3979, 3980) | Novel Protein sim. GBank gij5701727[db]BAA83074.1] - (AB024729) alpha-1,3-D-mannoside beta-1,4-N-acetylglucosaminyltransferase IV-homologue [Homo sapiens] | | | 264508, 264757, 264764, 18108381 |
| 1891 | 95353114 (3981, 3982) | Novel Protein sim. GBank gij4240287[db]BAA74922.1] - (AB020706) KIAA0899 protein [Homo sapiens] | Contains protein domain (PF01602) - Adaplin N terminal region | glycoprotein | 18108394, 56182575, 22278994, 35696286, 56994075, 22278997, 22278999, 29331822, 29331824, 29331825, 60432289, 29331828, 264508, 264906, 264907, 264908, 56182435, 264510, 265007, 21906754, 33109954, 87168474, 265017, 265018, 265019, 264762, 18108351, 264763, 264683, 264369, 264286, 264685, 264766, 264687, 264769, 21906765, 21906768, 21906769, 55811957, 265020, 60431528, 263974, 18108379, 35695855, 264555, 264557, 264639, 83373044, 18108384, 87168518, 60432113, 22279000, 22279002, 264564, 264486 |
| 1892 | 95317232 (3983, 3984) | Novel Protein sim. GBank gij2246532 (U93872) - ORF 73, contains large complex repeat CR 73 [Kaposi's sarcoma-associated herpesvirus] | Contains protein domain (PF00096) - Zinc finger, C2H2 type | struc | 28331827, 264906, 264907, 264909, 265007, 264603, 264766, 264688, 264768, 21906768, 264628, 264635, 264636, 18108385, 56526486, 264566, 264567 |
| 1893 | 80054763 (3985, 3986) | Novel Protein sim. GBank gij2565091 (U80761) - CTG26 alternate open reading frame [Homo sapiens] | | UNCLASSIFIED | 264592, 35696423 |

| | | | | | |
|------|-----------------------|---|---|--------------|---|
| 1984 | 94328114 (3987, 3988) | Novel Protein sim. GBank gi15630077 gb AAD45822.1 AC00601 - (AC006017) similar to ALR; similar to AAC51735 (PID:g2358287) [Homo sapiens] | Contains protein domain (PF00856) - SET domain | mapolymerase | 264488, 22278997, 22278998, 22278999, 29331822, 29331824, 29331825, 29331826, 29331827, 29331828, 264905, 264907, 264908, 265008, 60432229, 33657402, 60433356, 264757, 60433438, 264758, 33109954, 265011, 265017, 265018, 265019, 264684, 264369, 264685, 264686, 264768, 21906765, 21906767, 21906768, 21906769, 265020, 264690, 18108362, 264693, 85274620, 18108370, 264635, 264555, 264556, 264557, 56182323, 83373044, 56526486, 22279000, 22279002, 264564 264488, 18108396, 22278994, 56984075, 22278996, 35696286, 22278997, 22278999, 264259, 29147620, 56182181, 29331824, 60432289, 29331826, 29331827, 35696052, 29146499, 264905, 264907, 66712502, 56182435, 265006, 265007, 265008, 265009, 60431735, 60433356, 33657402, 264595, 55812038, 33657084, 55811386, 85658542, 265010, 265011, 265017, 265018, 265019, 264761, 264762, 264448, 264683, 264764, 264288, 264766, 264686, 264768, 264769, 56181562, 264689, 21906765, 21906768, 21906767, 29148627, 21906768, 21906769, 29148629, 29148784, 265020, 265021, 264690, 18108361, 264693, 27486262, 27486264, 27486265, 18108370, 60431528, 18108374, 18108377, 35696423, 55811578, 65274791, 35695855, 264631, 264634, 264635, 264555, 264636, 60431850, 264557, 264558, 264559, 83373044, 20798451, 87186518, 264404, 60432113, 264567 264564 |
| 1995 | 95414353 (3989, 3990) | Novel Protein sim. GBank gi4827040 ref NP_005110.1 pTRAP - thyroid hormone receptor-associated protein, 150 kDa subunit | | UNCLASSIFIED | |
| 1996 | 80254186 (3991, 3992) | Novel Protein sim. GBank gi1791146 emb CAA60020 - (XB6028) extensin-like protein [Vigna unguiculata] | | UNCLASSIFIED | |
| 1997 | 87028423 (3993, 3994) | Novel Protein sim. GBank gi12642034 (AF034547) - protein phosphatase M130 myosin binding subunit [Ovis aries] | Contains protein domain (PF00023) - Ank repeat | phosphatase | 264908, 264909, 264592, 264593 |
| 1998 | 85262704 (3995, 3996) | Novel Protein sim. GBank gi14589634 dbj BAA76839.1 - (AB023212) KIAA0995 protein [Homo sapiens] | | UNCLASSIFIED | 264113, 264685, 264555, 264567 |

| | | | | | |
|------|-----------------------|--|--|--------------|---|
| 1999 | 94324903 (3997, 3998) | Novel Protein sim. GBank gij5225312igb/AAD40846.1(AF072441) calchectin binding protein cabin 1 [Homo sapiens] | Contains protein domain (PF00515) - TPR Domain | UNCLASSIFIED | 18108394, 18108397, 35696286, 50424269, 29331827, 29331828, 35696052, 265006, 264512, 55811388, 265010, 265018, 265019, 5581150, 18108351, 264763, 264682, 264369, 264685, 264686, 56181582, 265020, 264691, 33657023, 264693, 33657109, 27486264, 18108370, 18108379, 35695855, 264634, 264635, 264636, 264555, 264557, 56182323, 18108382, 264559, 83373044, 60432113, 22279000, 264563, 264564, 264566 |
| 2000 | 95413705 (3999, 4000) | Novel Protein sim. GBank gij1723232ispjQ10155YATA_SCHPO - HYPOTHETICAL 90.6 KD PROTEIN C1D4.10 IN CHROMOSOME 1 | | UNCLASSIFIED | 52646365, 52646842, 22278994, 22278995, 22278998, 35696286, 22278997, 22278998, 22278999, 264259, 52645080, 29331822, 29331828, 29331827, 29331828, 35696052, 264106, 29331830, 52644045, 265007, 265008, 60170831, 264592, 264593, 33657402, 60433438, 21906754, 52644296, 265017, 265018, 265019, 264761, 264369, 264288, 52644229, 21906765, 21906766, 21906767, 21906768, 21906769, 35695917, 265020, 265021, 60170815, 52644150, 33657023, 65274620, 52645129, 27486261, 27486262, 27486264, 35695763, 35696423, 35695855, 264631, 52644332, 56182323, 60170394, 83373044, 56526488, 22279002, 264566, 264567 |
| 2001 | 95072534 (4001, 4002) | Novel Protein sim. GBank gij107560[pri]B38637 - Ras inhibitor (clone JC265) - human (fragment) | | UNCLASSIFIED | 264769, 52644229, 65274572, 21906768, 22278996, 35696286, 35695917, 265020, 22278999, 264534, 264490, 264259, 264692, 60432289, 33657109, 35696052, 264508, 264509, 18108370, 60431528, 18108374, 35696423, 65274791, 35695855, 264510, 264511, 264512, 265009, 264634, 264636, 264555, 264556, 264638, 264557, 264558, 264559, 60433438, 83373044, 264759, 18108385, 265011, 264600, 264601, 60432113, 264603, 264604, 264605, 264448, 264288, 264765 |
| 2002 | 80236368 (4003, 4004) | Novel Protein sim. GBank gij729433ispjP38637IER60_BOVIN - PROBABLE PROTEIN DISULFIDE ISOMERASE ER-60 PRECURSOR (ERP60) (58 KD MICROSOMAL PROTEIN) (P58) | Contains protein domain (PF00085) - Thioredoxin | isomerase | 264907, 265006, 264910, 264603, 264692, 264629, 18108374, 264556, 264557 |
| 2003 | 80074449 (4005, 4006) | Novel Protein sim. GBank gij86388[pri]A27040 - neurofilament triplet M protein - chicken (fragment) | | UNCLASSIFIED | 264905, 264906, 264908, 264910, 264596, 265017, 18108351, 264692, 264629, 264634, 264565 |

| | | | | | |
|------|-----------------------|---|---|--------------|--|
| 2004 | 95317318 (4007, 4008) | Novel Protein sim. GBank gj4884249[embjCAB43230.1] - (AL048996) hypothetical protein [Homo sapiens] | Contains protein domain (PF00076) - RNA recognition motif. (a.k.a. RRM, RBD, or RNP domain) | dna_rna_bind | 52645156, 52646842, 52646365, 56182575, 22278994, 22278995, 56994075, 22278996, 35686286, 22278997, 22278998, 22278999, 60432049, 264259, 52645080, 29331824, 29331826, 29331827, 35686052, 29331828, 33656970, 29331830, 264908, 264592, 60433356, 33657402, 52646317, 21906754, 33657084, 52644296, 87168474, 87168559, 265017, 265018, 265019, 264763, 264683, 264288, 52644229, 21906768, 21906769, 35695917, 21906767, 21906768, 21906769, 35695917, 265020, 265021, 265022, 52644150, 33657023, 52645129, 33657109, 33657182, 27486261, 27486282, 33657349, 27486265, 35695763, 18108374, 18108376, 18108377, 35696423, 35695855, 264631, 52644332, 264558, 18108385, 56526486, 87168518, 60432113, 264483, 264488, 264906 |
| 2005 | 87400864 (4009, 4010) | Novel Protein sim. GBank gj3879501[embjCAA87795] - (Z47812) similar to ubiquitin carboxyl-terminal hydrolase; cDNA EST EMBL:D33368 comes from this gene; cDNA EST EMBL:D33965 comes from this gene; cDNA EST EMBL:D33822 comes from this gene; cDNA EST EMBL:D34547 comes from this ge... | | ubiquitin | |
| 2006 | 95351177 (4011, 4012) | Novel Protein sim. GBank gj4106673[embjCAA22613] - (AL035064) queuine tRNA-ribosyltransferase [Schizosaccharomyces pombe] | Contains protein domain (PF01702) - Queuine tRNA-ribosyltransferase | UNCLASSIFIED | 56182575, 56181686, 22278995, 22278996, 22278997, 22278998, 22278999, 264259, 56182181, 60424269, 66714117, 35696052, 264906, 66712502, 264909, 264510, 60433356, 85658542, 265010, 265018, 265019, 264682, 264448, 264288, 264768, 29148627, 21906769, 29148784, 35695917, 60170615, 264691, 33657023, 65274620, 33657109, 55810764, 55811576, 35695855, 87168518, 60432113, 264563, 264482 |
| 2007 | 94325556 (4013, 4014) | Novel Protein sim. GBank gj2682161[dbjBAA23712] - (AB007900) HH0452 cDNA clone for KIAA0440 has a 438-bp insertion at position 1711 of the sequence of KIAA0440. [Homo sapiens] | | UNCLASSIFIED | 264488, 263994, 35696052, 264508, 264905, 264906, 264907, 264908, 264909, 265007, 264910, 264592, 264595, 264758, 265011, 264760, 264762, 264764, 264766, 264685, 264767, 264768, 264769, 55811957, 35695917, 265020, 264691, 264693, 264628, 264629, 65274791, 35695855, 264631, 264632, 264634, 264635, 264637, 264638, 264639, 264566 |
| 2008 | 85084428 (4015, 4016) | Novel Protein sim. GBank gj1550783[embjCAA69257] - (Y07960) homeodomain protein [Mus musculus] | Contains protein domain (PF00046) - Homeobox domain | homeobox | 264909, 264768, 35695855 |

| | | | | |
|------|-----------------------|--|--|--|
| 2008 | 85748240 (4017, 4016) | Novel Protein sim. GBank gi 3882305 dbj BAA34512.1 - (AB018335) KIAA0792 protein [Homo sapiens] | UNCLASSIFIED | 22278998, 264259, 264810, 264591, 265017, 264881, 264683, 21906768, 264691, 33657182, 33657349, 264631, 87168518, 264404, 22279002, 264563 |
| 2010 | 85422458 (4019, 4020) | Novel Protein sim. GBank gi 5262629 emb CAB45753.1 - (AL080164) hypothetical protein [Homo sapiens] | Contains protein domain (PF00057) - eph Low-density lipoprotein receptor domain class A | 52644507, 52645156, 52646365, 52646842, 18108397, 65274572, 22278994, 58994075, 35696286, 22278996, 22278997, 22278999, 264259, 29331822, 52645080, 29331824, 29331825, 29331826, 29331827, 29331828, 264511, 265007, 264512, 265008, 265009, 60432229, 60433356, 21906754, 52646317, 33109954, 52644296, 87168474, 87168559, 265017, 265018, 265019, 264681, 264685, 264687, 52644229, 264689, 21906765, 21906766, 21906767, 21906768, 35695917, 265020, 52644150, 264691, 264692, 33657023, 263987, 52645129, 35695763, 18108376, 35696423, 65274791, 35695855, 264631, 264634, 60431850, 264637, 264638, 52644332, 60170394, 18108385, 87168518, 22279002, 264564, 264565, 264566, 264567 |
| 2011 | 94328149 (4021, 4022) | Novel Protein sim. GBank gi 3347953 (AF076183) - cytosolic sorting protein PACS-1a [Rattus norvegicus] | UNCLASSIFIED | 56182575, 56994075, 22278999, 264259, 29331824, 29331826, 29331827, 29331828, 35696052, 264906, 66712502, 265006, 265007, 265008, 265010, 265011, 265017, 265019, 264681, 264448, 264683, 264369, 264288, 264685, 264766, 264687, 21906765, 21906767, 21906768, 21906769, 265020, 265022, 264691, 33657023, 65274620, 33657109, 264629, 264557, 264559, 83373044, 87168518, 60432113, 22279002 |
| 2012 | 87772137 (4023, 4024) | Novel Protein sim. GBank gi 1086678 (U41020) - coded for by C. elegans cDNA yk100g4.5; coded for by C. elegans cDNA yk100g4.3; weakly similar to human SREBP-2 basic-helix-loop-helix-leucine zipper transcription factor [Caenorhabditis elegans] | Contains protein domain (PF00409) - Kinesin light chain repeat | 264259, 29331822, 29331824, 29331825, 29146498, 87168559, 265019, 264682, 264288, 264686, 21906764, 265020, 265022, 33657023, 264693, 33657109, 55811576, 264632, 264558, 56182323, 264639, 18108383, 18108384, 18108388, 22279000, 22279002, 264567 |
| 2013 | 94843842 (4025, 4026) | Novel Protein sim. GBank gi 4507985 ref NP_003427.1 pZNF1 - zinc finger protein 135 (clone pTZ-17) | Contains protein domain (PF00096) - Zinc finger, C2H2 type | 18108398, 264908, 265007, 265010, 265018, 265019, 264689, 21906767, 265020, 264692 |
| 2014 | 87347840 (4027, 4028) | Novel Protein sim. GBank gi 127720 sp P20938 MYPO_HETFR - MYELIN P0 PROTEIN PRECURSOR | UNCLASSIFIED | 264488, 29331826, 264807, 264636, 264555, 264639, 264558 |

| | | | | | |
|------|-----------------------|---|---|--------------|---|
| 2015 | 88094922 (4029, 4030) | Novel Protein sim. GBank gij181286 pir S22697 - extensin - Volvox carteri (fragment) | | UNCLASSIFIED | 56182575, 35696286, 264259, 35696052, 264508, 264906, 264907, 264510, 264512, 87168474, 265010, 264681, 264288, 264689, 264628, 35698423, 35698555, 264639, 264563, 264564 |
| 2016 | 85298641 (4031, 4032) | Novel Protein sim. GBank gij285046 pir S26413 - t-complex protein Tcp-10 - mouse | | struct | 264102, 264508, 264110, 265009, 33108954, 21906768, 265021, 33657109, 27486282, 263972, 18108374, 263976, 264555, 264564 |
| 2017 | 79464293 (4033, 4034) | Novel Protein sim. GBank gij124735 sp P18175 INVO_PIG - INVOLUCRIN | | UNCLASSIFIED | 264685, 264636 |
| 2018 | 79637067 (4035, 4036) | Novel Protein sim. GBank gij21439 0 pir S68216 - phosphatase-1 glycogen-binding (GL)-chain - rat | | phosphatase | 264107, 264110, 264112, 265017, 263976 |
| 2019 | 87787800 (4037, 4038) | Novel Protein sim. GBank gij2078483 (U43200) - antifreeze glycopeptide AFGP polyprotein precursor [Boreogadus saldai] | | UNCLASSIFIED | 264259, 264508, 264591, 265018, 264682, 264288, 264688, 22278002 |
| 2020 | 84674476 (4039, 4040) | Novel Protein sim. GBank gij585084 sp Q07803 EFGM_RAT - ELONGATION FACTOR G, MITOCHONDRIAL PRECURSOR (MEF-G) | | | 56994075, 264593, 33109954, 21908754, 21908768, 33657023, 33657109, 27486261, 87168518 |
| 2021 | 86718818 (4041, 4042) | Novel Protein sim. GBank gij4218005 (AC006135) - putative vicilin storage protein (globulin-like) [Arabidopsis thaliana] | | | 264757, 264767, 60170615, 18108385 |
| 2022 | 95295665 (4043, 4044) | Novel Protein sim. GBank gij5410230 gb AAD2982.1 AF07334 - (AF073344) ubiquitin specific protease 3 [Homo sapiens] | | | 18108394, 22278999, 264259, 264905, 264906, 264908, 264595, 264762, 264769, 264634, 264636, 87168518, 60432113, 22279000, 264482, 264565 |
| 2023 | 87722976 (4045, 4046) | Novel Protein sim. GBank gij5410230 gb AAD2982.1 AF07334 - (AF073344) ubiquitin specific protease 3 [Homo sapiens] | Contains protein domain (PF00442) - Ubiquitin carboxyl-terminal hydrolases family 2 | ubiquitin | 60433438, 265017, 264686, 264692, 264693, 264636 |
| 2024 | 87896443 (4047, 4048) | Novel Protein sim. GBank gij4755188 gb AAD29055.1 AC00701 - (AC007018) unknown protein [Arabidopsis thaliana] | Contains protein domain (PF00637) - 7-fold repeat in Clathrin and VPS | UNCLASSIFIED | 22278997, 264509, 264906, 264908, 55812038, 265017, 265021, 265022, 60170615, 264556 |
| 2025 | 87858863 (4049, 4050) | Novel Protein sim. GBank gij1655899 emb CAA69032 - (Y07752) perophorin-5 [Volvox carteri] | | UNCLASSIFIED | 56994075, 60432049, 264508, 66712502, 264112, 60170831, 87168559, 264288, 264688, 264689, 21906766, 33657109, 18108370, 264638, 18108385, 60432113, 22279000, 22279002, 264564, 264566, 264567 |
| 2026 | 84122114 (4051, 4052) | Novel Protein sim. GBank gij1655899 emb CAA69032 - (Y07752) perophorin-5 [Volvox carteri] | | | 263978, 264634, 264486 |
| 2027 | 80249001 (4053, 4054) | | | UNCLASSIFIED | |

| | | | | | |
|------|-----------------------|--|--|-------------------|--|
| 2028 | 94699884 (4055, 4056) | Novel Protein sim. GBank gi 1330345 (U58755) - coded for by C. elegans cDNA yk34b1.5; coded for by C. elegans cDNA yk13n10.5; coded for by C. elegans cDNA yk46e8.5; coded for by C. elegans cDNA yk46d5.5; coded for by C. elegans cDNA yk43c2.5; coded for by C. elegans cDNA yk46e8.... | Contains protein domain (PF00632) - HECT-domain (ubiquitin-transferase). | ubiquitin | 52644507, 52645156, 52646842, 56182575, 56994075, 35696286, 22278997, 22278998, 22278999, 60432049, 264259, 52645080, 29331822, 29331824, 66714117, 29331825, 29331828, 29331827, 29331828, 35696052, 264906, 264907, 264908, 29331830, 52644045, 56182435, 265006, 265009, 60432228, 33657402, 264595, 264757, 55812038, 21906754, 52646317, 52644296, 265010, 265011, 87168559, 265017, 265018, 265019, 264448, 18108354, 264288, 264369, 264766, 52644228, 21906765, 21906766, 21906767, 21906768, 21906769, 55811957, 35695917, 265020, 265021, 265022, 52644150, 33657023, 65274820, 33657109, 52645129, 18108368, 27486261, 27486262, 27486264, 27486265, 35695763, 264629, 55811578, 35698423, 35695855, 264635, 264636, 52644332, 264558, 83373044, 56526486, 22279000, 22279002, 264563, 265009, 264595, 85658542, 264555, 264556, 264557, 264558, 264559, 83373044 |
| 2029 | 85362032 (4057, 4058) | Novel Protein sim. GBank gi 3599940 (AF017368) - facio-genital dysplasia protein 2 [Mus musculus] | Contains protein domain (PF00621) - RhoGEF domain | UNCLASSIFIED | 18108394, 56994075, 22278997, 22278999, 264259, 29331822, 29331824, 29331825, 66714117, 60432289, 29331826, 264108, 66712502, 264828, 265009, 265018, 265019, 264681, 264682, 264684, 264685, 56181562, 264689, 21906769, 265022, 264682, 264693, 264628, 18108370, 18108374, 264634, 264636, 264556, 18108385, 87168518, 22279002, 264565 |
| 2030 | 91213734 (4059, 4060) | Novel Protein sim. GBank gi 5630080 gb AAD45825.1 AC00489 - (AC004890) similar to HUB1; similar to BAA24380 (PID:g2789430) [Homo sapiens] | Contains protein domain (PF00096) - Zinc finger, C2H2 type | transcript factor | 264591, 55811957, 18108365, 264557, 264558, 18108382, 18108384 |
| 2031 | 80245281 (4061, 4062) | Novel Protein sim. GBank gi 5689491 dbj BAA83029.1 - (AB029000) KIAA1077 protein [Homo sapiens] | Contains protein domain (PF00884) - Sulfatase | hydrolase | 55274572, 35696286, 29331824, 264908, 265009, 264593, 265018, 264288, 264686, 264769, 21906768, 21906767, 29148627, 264628, 35696423, 264634, 264556, 18108381, 60170394, 264559, 83373044, 18108385, 264482, 264484 |
| 2032 | 91232607 (4063, 4064) | Novel Protein sim. GBank gi 2494828 sp Q64686 CAG7_RAT - ALPHA-N-ACETYLGLACTOSAMINIDE ALPHA-2-6-SIALYLTRANSFERASE (ST6GALNACIII) (STY) | Contains protein domain (PF00076) - RNA recognition motif (a.k.a. RRM, RBD, or RNP domain) | synthase | 56181562, 264628, 264632, 264555, 264556 |
| 2033 | 95000809 (4065, 4066) | Novel Protein sim. GBank gi 4826984 ref NP_005147.1 PROD1 - UNKNOWN | | | |
| 2034 | 91232529 (4067, 4068) | Novel Protein sim. GBank gi 4826984 ref NP_005147.1 PROD1 - UNKNOWN | | | |

| | | | | | |
|------|-----------------------|--|---|--------------|--|
| 2035 | 83553451 (4069, 4070) | | | | 264369, 264686, 265022, 56526486, 264567 |
| 2036 | 87115833 (4071, 4072) | | | | 29331827, 29331828, 264682, 264369, 29148627, 60432113 |
| 2037 | 94324833 (4073, 4074) | Novel Protein sim. GBank gi 2734081 (AF000195) - similar to oxysterol-binding proteins [Caenorhabditis elegans] | | UNCLASSIFIED | 65274572, 22278995, 22278996, 56994075, 35696286, 22278997, 22278998, 22278999, 264259, 29331824, 60432289, 29331826, 29331828, 35696052, 264907, 29331830, 66712502, 56182435, 265008, 265009, 60170831, 264584, 55812038, 33109954, 21906754, 87168559, 265017, 265018, 265019, 264762, 264369, 264288, 21906765, 21906767, 21906768, 21906769, 55811957, 35695917, 265020, 265021, 265022, 52644150, 33657023, 33657109, 33657182, 35695763, 35695855, 264632, 264634, 264636, 56182323, 83373044, 60432113, 22279000, 22279002, 264563 |
| 2038 | 95422384 (4075, 4076) | Novel Protein sim. GBank gi 3880625 emb CAB07859 - (Z93785) predicted using GeneFinder; similar to RNA recognition motif (aka RRM, RBD, or RNP domain); cDNA EST EMBL: T01682 comes from this gene; cDNA EST EMBL: M75823 comes from this gene; cDNA EST EMBL: D27559 comes from this ge... | Contains protein domain (PF01412) - Putative GTP-ase activating protein (for Arf) | UNCLASSIFIED | 22278995, 22278996, 56994075, 264259, 29331824, 35696052, 264905, 264906, 52644045, 265007, 265009, 87168559, 265017, 18108351, 264448, 264389, 264768, 264767, 264686, 18108358, 21906765, 21906769, 52644150, 33657023, 264692, 18108362, 33657109, 27486262, 18108370, 18108374, 18108379, 35696423, 65274791, 264632, 264636, 18108383, 83373044, 18108385, 87168518, 22279000, 22279002, 264563, 264584, 264566 |
| 2039 | 85514626 (4077, 4078) | Novel Protein sim. GBank gi 2224653 dbj BAA20813 - (AB002354) KIAA0356 [Homo sapiens] | Contains protein domain (PF00097) - Zinc finger, C3HC4 type (RING finger) | UNCLASSIFIED | 22278997, 264259, 29331822, 264905, 264906, 264907, 264908, 264909, 264510, 265009, 264910, 264593, 264758, 265011, 265018, 264762, 264288, 264766, 264768, 264769, 21906766, 33657023, 264692, 264693, 33657109, 35696423, 264631, 264632, 264634, 264635, 264636, 264637, 264639, 87168518, 264486 |
| 2040 | 95308417 (4079, 4080) | Novel Protein sim. GBank | | UNCLASSIFIED | 264592 |
| 2041 | 95071736 (4081, 4082) | gi 2500825 sp P700JRP2 MOUSE - DNA-DIRECTED RNA POLYMERASE I 135 KD POLYPEPTIDE (RNA POLYMERASE I SUBUNIT 2) (RPA135) | mapolymerase | | 264488, 22278998, 35696052, 264805, 264907, 264908, 264910, 265018, 264605, 265019, 18108351, 264766, 264769, 21906766, 265021, 265022, 264692, 33657109, 264628, 264629, 35696423, 35695855, 264637, 264638, 264563, 264564, 264565, 264567 |

| | | | | | |
|------|-----------------------|--|--|--------------|--|
| 2042 | 95307447 (4083, 4084) | Novel Protein sim. GBank gi 4406590 gb AAD20040 - (AF131766) Similar to Ena-VASP like protein [Homo sapiens] | Contains protein domain (PF00568) - WH1 domain | UNCLASSIFIED | 60424179, 35696286, 264259, 29331826, 35696052, 29331828, 264508, 264509, 264907, 264909, 264510, 264511, 265009, 264910, 264591, 60433356, 264595, 265017, 265019, 264681, 264764, 264369, 264765, 264684, 264288, 264766, 264686, 52644229, 264769, 21906765, 35695917, 264535, 52644150, 264691, 264692, 18108365, 27486281, 27486282, 27486285, 18108374, 35696423, 65274791, 35695855, 264555, 264558, 60170394, 18108385, 264404, 22279000, 22279002, 264482, 264583, 264564, 264566 |
| 2043 | 94328076 (4085, 4086) | Novel Protein sim. GBank gi 5052554 gb AAD38607.1 AF145632) BCDNA GH06032 [Drosophila melanogaster] | Contains protein domain (PF00122) - E1-E2 ATPase | transport | 264488, 52644507, 52646365, 56694075, 22278997, 22278999, 20281171, 264259, 29331822, 29331824, 66714117, 29331826, 29331828, 33656970, 29146498, 264509, 264908, 52644045, 58182435, 265006, 33657402, 21906754, 52644296, 87168559, 265017, 265018, 265019, 264681, 264288, 264766, 264685, 264686, 21906766, 21906767, 21906768, 21906769, 265020, 265021, 60170815, 264691, 33657023, 264693, 65274620, 33657109, 33657182, 27486261, 27486262, 33657349, 35695763, 18108374, 55811576, 35695855, 18108380, 18108381, 60170394, 55182323, 264558, 83373044, 18108385, 56526486, 87168518, 60432113, 22279000, 264567 |
| 2044 | 87106927 (4087, 4088) | Novel Protein sim. GBank gi 2246532 (U93872) - ORF 73, contains large complex repeat CR 73 [Kaposi's sarcoma-associated herpesvirus] | | situct | 264093, 29331827, 264905, 66712502, 264592, 264889, 21908765, 21908769, 265020, 264692, 264482, 264566 |
| 2045 | 79635532 (4089, 4090) | Novel Protein sim. GBank gi 4406688 gb AAD20062 - (AF131852) Unknown [Homo sapiens] | | | 264692 |
| 2046 | 87320849 (4091, 4092) | Novel Protein sim. GBank gi 4406688 gb AAD20062 - (AF131852) Unknown [Homo sapiens] | | | 264259, 264906, 264683, 22279002 |
| 2047 | 84578801 (4093, 4094) | Novel Protein sim. GBank gi 4101720 (AF006466) - lymphocyte specific formin related protein [Mus musculus] | | | 22278999, 29147620, 29331824, 29146498, 264508, 265007, 265008, 265019, 264605, 264681, 29148627, 29148629, 265021, 33657023, 18108365, 33657109, 33657182, 18108377, 264556, 264638, 264559, 18108388 |
| 2048 | 84606378 (4095, 4096) | Novel Protein sim. GBank gi 4589656 dbj BAA76850.1 - (AB023223) KIAA1006 protein [Homo sapiens] | | UNCLASSIFIED | 264909 |
| 2049 | 88094680 (4097, 4098) | Novel Protein sim. GBank gi 4589656 dbj BAA76850.1 - (AB023223) KIAA1006 protein [Homo sapiens] | | UNCLASSIFIED | 264488, 264259, 29331824, 29331828, 35696052, 264906, 264907, 264908, 264909, 264910, 264603, 264763, 21906767, 21906768, 264629, 264634, 264637, 22279002, 264564, 264565, 264566, 264567 |

| | | | | | | | |
|------|-----------------------|--|--|--|--|--------------|---|
| 2050 | 79633835 (4099, 4100) | | | | | UNCLASSIFIED | 264693 |
| 2051 | 87780168 (4101, 4102) | | | | | UNCLASSIFIED | 264488, 264259, 264509, 264906, 264907, 264769, 18108374, 35698423, 264563, 264586, 264486 |
| 2052 | 88096393 (4103, 4104) | Novel Protein sim. GBank gj14529889[gblAAD21812.1] - (AF134726) G9A [Homo sapiens] | Contains protein domain (PF00856) - SET domain | | | kinase | 264488, 263994, 35696052, 264508, 264905, 264509, 264908, 264907, 264909, 264113, 264511, 285009, 264910, 60170831, 264592, 264758, 265010, 265011, 264605, 264760, 264682, 264764, 264369, 264766, 264686, 264768, 264769, 52644229, 264689, 35695917, 33657023, 33657109, 264628, 18108374, 35696423, 55811576, 35695855, 264630, 264631, 264632, 264634, 264635, 264636, 264556, 264638, 264639, 18108385, 56526486, 60432113, 264563, 264564, 264566, 264486, 264567, 264488, 263994, 35696052, 264508, 264905, 264509, 264908, 264907, 264908, 264909, 264113, 264511, 265009, 264910, 60170831, 264592, 264758, 265010, 265011, 264605, 264760, 264682, 264764, 264369, 264766, 264686, 264768, 264769, 52644229, 264689, 35695917, 33657023, 33657109, 264628, 18108374, 35698423, 55811576, 35695855, 264630, 264631, 264632, 264634, 264635, 264636, 264556, 264638, 264639, 18108385, 56526486, 60432113, 264563, 264564, 264566, 264486, 264567, 264488, 263994, 264769, 52644229, 264689, 35695917, 33657023, 33657109, 264628, 18108374, 35698423, 55811576, 35695855, 264630, 264631, 264632, 264634, 264635, 264636, 264556, 264638, 264639, 18108385, 56526486, 60432113, 264563, 264564, 264566, 264486, 264567, 264488, 263994, 264769, 52644229, 264689, 35695917, 33657023, 33657109, 264628, 18108374, 35698423, 55811576, 35695855, 264630, 264631, 264632, 264634, 264635, 264636, 264556, 264638, 264639, 18108385, 56526486, 60432113, 264563, 264564, 264566, 264486, 264567, 264488, 263994, 264769, 52644229, 264689, 35695917, 33657023, 33657109, 264628, 18108374, 35698423, 55811576, 35695855, 264630, 264631, 264632, 264634, 264635, 264636, 264556, 264638, 264639, 18108385, 56526486, 60432113, 264563, 264564, 264566, 264486, 264567, 264488, 263994, 264769, 52644229, 264689, 35695917, 33657023, 33657109, 264628, 18108374, 35698423, 55811576, 35695855, 264630, 264631, 264632, 264634, 264635, 264636, 264556, 264638, 264639, 18108385, 56526486, 60432113, 264563, 264564, 264566, 264486, 264567, 264488, 263994, 264769, 52644229, 264689, 35695917, 33657023, 33657109, 264628, 18108374, 35698423, 55811576, 35695855, 264630, 264631, 264632, 264634, 264635, 264636, 264556, 264638, 264639, 18108385, 56526486, 60432113, 264563, 264564, 264566, 264486, 264567, 264488, 263994, 264769, 52644229, 264689, 35695917, 33657023, 33657109, 264628, 18108374, 35698423, 55811576, 35695855, 264630, 264631, 264632, 264634, 264635, 264636, 264556, 264638, 264639, 18108385, 56526486, 60432113, 264563, 264564, 264566, 264486, 264567, 264488, 263994, 264769, 52644229, 264689, 35695917, 33657023, 33657109, 264628, 18108374, 35698423, 55811576, 35695855, 264630, 264631, 264632, 264634, 264635, 264636, 264556, 264638, 264639, 18108385, 56526486, 60432113, 264563, 264564, 264566, 264486, 264567, 264488, 263994, 264769, 52644229, 264689, 35695917, 33657023, 33657109, 264628, 18108374, 35698423, 55811576, 35695855, 264630, 264631, 264632, 264634, 264635, 264636, 264556, 264638, 264639, 18108385, 56526486, 60432113, 264563, 264564, 264566, 264486, 264567, 264488, 263994, 264769, 52644229, 264689, 35695917, 33657023, 33657109, 264628, 18108374, 35698423, 55811576, 35695855, 264630, 264631, 264632, 264634, 264635, 264636, 264556, 264638, 264639, 18108385, 56526486, 60432113, 264563, 264564, 264566, 264486, 264567, 264488, 263994, 264769, 52644229, 264689, 35695917, 33657023, 33657109, 264628, 18108374, 35698423, 55811576, 35695855, 264630, 264631, 264632, 264634, 264635, 264636, 264556, 264638, 264639, 18108385, 56526486, 60432113, 264563, 264564, 264566, 264486, 264567, 264488, 263994, 264769, 52644229, 264689, 35695917, 33657023, 33657109, 264628, 18108374, 35698423, 55811576, 35695855, 264630, 264631, 264632, 264634, 264635, 264636, 264556, 264638, 264639, |

| | | | | | |
|------|-----------------------|---|---|--------------|--|
| 2056 | 86177396 (4111, 4112) | Novel Protein sim. GBank glj4828960refNP_005042:lpQARS - glutamine-IRNA synthetase | Contains protein domain (PF00749) - IRNA synthetases class I (E and Q) | - synthase | 264488, 52645156, 56182575, 22278994, 35696286, 56994075, 22278996, 22278998, 22278999, 60432049, 264259, 29331824, 60432289, 29331827, 29331828, 33656970, 284104, 264908, 264908, 265006, 285008, 60170831, 264591, 60432229, 60433438, 18108348, 21908754, 33657084, 52644296, 87168474, 265010, 87168559, 265017, 265018, 264760, 18108351, 264681, 264682, 264448, 264683, 264369, 264288, 264685, 264687, 264688, 264689, 21908765, 21908766, 21908767, 21908769, 55811957, 35695917, 265022, 33657023, 18108362, 33657109, 18108368, 33657182, 27486261, 27486264, 27486265, 33657349, 264628, 18108370, 264629, 18108374, 18108377, 18108379, 35696423, 55811576, 20281152, 264636, 264952, 18108385, 18108388, 87168518, 264482, 264565, 264586, 264567 |
| 2057 | 87877905 (4113, 4114) | Novel Protein sim. GBank glj728850spIP08640JAMYH_YEAST - GLUCOAMYLASE S1/S2 PRECURSOR (GLUCAN 1,4-ALPHA- GLUCOSIDASE) (1,4-ALPHA-D-GLUCAN GLUCOHYDROLASE) | | UNCLASSIFIED | 52646842, 52646365, 56182575, 35696286, 22278996, 22278997, 22278998, 264093, 52645080, 35696052, 29331828, 33656970, 265009, 52646317, 55811386, 52644296, 52644229, 21908769, 35695917, 265021, 60170615, 52644150, 33657109, 33657182, 27486261, 27486262, 35695763, 35696423, 35695855, 52644332 |
| 2058 | 86276896 (4115, 4116) | | | | 265007, 265008, 264591 |
| 2059 | 79866684 (4117, 4118) | Novel Protein sim. GBank glj119714spIP13983JEXTN_TOBAC - EXTENSIN PRECURSOR (CELL WALL HYDROXYPROLINE-RICH GLYCOPROTEIN) | | UNCLASSIFIED | 29331825, 264682, 264686, 264691, 264693, 22279002 |
| 2060 | 83050800 (4119, 4120) | Novel Protein sim. GBank glj2811122 (U87318) - NaDC-2 [Xenopus laevis] | | UNCLASSIFIED | 56182575, 29331824, 29331826, 264910, 55811957, 18108370, 55811576 |

| | | | | | |
|------|-----------------------|---|---|--------------|---|
| 2061 | 95362204 (4121, 4122) | Novel Protein sim. GBank gi 2496947 sp Q09298 YQ09_CAEEL - HYPOTHETICAL 141.2 KD PROTEIN EEED8.9 IN CHROMOSOME II | Contains protein domain (PF00069) - Eukaryotic protein kinase domain | kinase | 22278997, 22278999, 264259, 29331822, 29331824, 29331826, 29331827, 35696052, 29331828, 264906, 66712502, 29331830, 264908, 264909, 264112, 264511, 265007, 265009, 264910, 264591, 33657402, 21906754, 85658542, 265017, 265019, 264448, 264683, 264288, 264684, 264369, 264686, 264687, 18108358, 264689, 21906765, 21906767, 21906768, 21906769, 265020, 265022, 264691, 33657023, 33657109, 20281149, 18108379, 35695855, 264634, 264556, 264557, 264558, 18108382, 264559, 83373044, 18108384, 56528486, 60432113 |
| 2062 | 87028440 (4123, 4124) | Novel Protein sim. GBank gi 4502091 ref NP_001139.1 pANK2 - ankyrin 2, neuronal | Contains protein domain (PF00023) - Ank repeat | strut | 264905, 264628, 264907, 264629, 264908, 264909, 18108374, 263978, 35695855, 264512, 264635, 60431850, 264636, 264760, 264563, 18108351, 264762, 264565, 264764, 264487, 264766 |
| 2063 | 87601272 (4125, 4126) | Novel Protein sim. GBank gi 4589562 db BAA76803.1 - (AB023176) KIAA0959 protein [Homo sapiens] | Contains protein domain (PF00617) - RasGEF domain | oncogene | 22278994, 22278999, 264259, 29331827, 264906, 264909, 52644045, 264686, 21906767, 55811957, 264692, 18108365, 263972, 55811576, 18108384, 22279002, 264482, 264563, 264564, 264484 |
| 2064 | 95317253 (4127, 4128) | Novel Protein sim. GBank gi 1754515 db BAA13413.1 - (D87515) aminopeptidase-B [Rattus norvegicus] | | hydrolase | 264488, 52646365, 56994075, 35696286, 22278997, 22278998, 264259, 29331826, 60432289, 29331827, 29331828, 35696052, 264509, 265007, 265008, 60432229, 60433438, 21906754, 265010, 265011, 87168559, 265017, 265018, 264761, 18108351, 264682, 264369, 264288, 52644229, 21906765, 21906767, 21906768, 35695917, 33657109, 18108368, 18108374, 35696423, 35695855, 52644332, 264559, 60432113, 22279000, 22279002, 264566, 264486 |
| 2065 | 95092238 (4129, 4130) | Novel Protein sim. GBank gi 2507144 sp Q04205 TENS_CHICK - TENSIN | | kinase | 264569, 18108394, 56182181, 60432289, 29331826, 264905, 264906, 264908, 60431735, 60433356, 55811386, 85658542, 265018, 55811150, 264681, 264766, 264692, 60431528, 263974, 55810764, 35695855, 264631, 264634, 264635, 60431850, 264557, 83373044, 18108388, 22279000, 22279002 |
| 2066 | 85793402 (4131, 4132) | Novel Protein sim. GBank gi 160171 (M58295) - circumsporozoit protein [Plasmodium yoelii] | Contains protein domain (PF00096) - Zinc finger, C2H2 type | UNCLASSIFIED | 56182575, 264259, 264906, 264764, 264288, 56182323, 264567 |

| | | | | | |
|------|-----------------------|---|--|------------------|--|
| 2067 | 95303892 (4133, 4134) | | | | 35696286, 22278997, 22278998, 60432049, 264259, 60432289, 60433438, 264682, 264448, 264369, 264288, 18108355, 21906765, 21906768, 265022, 33657109, 35696423, 35695855, 264558, 264404, 264553, 264486 |
| 2068 | 84344754 (4135, 4136) | | | UNCLASSIFIED | 264687 |
| 2069 | 94319177 (4137, 4138) | Novel Protein sim. GBank gj 3152682 (AF064604) - KE03 protein [Homo sapiens] | Contains protein domain (PF00023) - Ank repeat | transcriptfactor | 60424179, 56182575, 22278995, 22278996, 56994075, 264259, 29331822, 29331824, 29331825, 35696052, 29331828, 33656870, 264509, 264905, 56182435, 265009, 60433358, 87168559, 265017, 265018, 264604, 265019, 264448, 264764, 264766, 21906765, 21906767, 21906768, 21906769, 265020, 265021, 33657023, 33657109, 263976, 264555, 264557, 56182323, 83373044, 87168518, 60432113, 22279000, 22279002 |
| 2070 | 85791380 (4139, 4140) | Novel Protein sim. GBank gj 5712131 gb AAD47379.1 AF12049 - (AF120499) DEM1 protein [Homo sapiens] | | UNCLASSIFIED | 35695917, 264905, 264628, 264908, 264638 |
| 2071 | 86946116 (4141, 4142) | Novel Protein sim. GBank gj 3551531 dbj BAA33016 - (AB017437) avena [Gallus gallus] | Contains protein domain (PF00568) - WH1 domain | | 18108398, 265006, 265007, 265008, 265009, 264594, 265010, 265011, 18108351, 18108354, 18108364, 18108365, 18108368, 264634, 18108381, 18108385, 18108386, 18108391 |
| 2072 | 91718428 (4143, 4144) | | Contains protein domain (PF00184) - Neurohypophyseal hormones, C-terminal Domain | | 22278995, 35696286, 22278997, 22278998, 22278999, 264490, 60432049, 264259, 29331822, 29331824, 29331826, 35696052, 265008, 33657402, 21906754, 265011, 265019, 18108351, 264682, 264369, 21906765, 21906766, 21906767, 21906768, 21906769, 35695917, 265020, 265021, 264680, 264692, 35696423, 264555, 264556, 264558, 22279000 |
| 2073 | 27825664 (4145, 4146) | Novel Protein sim. GBank gj 1504026 dbj BAA13212 - (D86976) similar to C.elegans protein (Z37093) [Homo sapiens] | | UNCLASSIFIED | 264556 |
| 2074 | 94324767 (4147, 4148) | Novel Protein sim. GBank gj 4240317 dbj BAA74937.1 - (AB020721) KIAA0914 protein [Homo sapiens] | | | 29331822, 264909, 264511, 265008, 264594, 264595, 265010, 265011, 265017, 265018, 265019, 264448, 264683, 265020, 265021, 18108370, 264632, 83373044, 264567 |

| | | | | | |
|------|-----------------------|---|--|--------------|--|
| 2075 | 84314886 (4149, 4150) | Novel Protein sim. GBank gij5138930 gb AAD40382.1 - (AF093880) transcription factor IIIB [Homo sapiens] | | | 18108394, 22278994, 22278996, 22278997, 22278998, 22278999, 264259, 29331822, 29331825, 29331827, 35696052, 29331828, 264905, 264907, 264908, 264510, 265007, 264910, 265009, 33657084, 264760, 264448, 264288, 264766, 264767, 264689, 21906766, 21906767, 21906769, 265021, 265022, 60170615, 33657023, 27486262, 27486265, 35696423, 35695855, 264631, 264634, 264639, 87168518, 22279002, 264563, 264486, 18108391 |
| 2076 | 87594118 (4151, 4152) | | | | 264259, 29331826, 264508, 264908, 264510, 265007, 265011, 264288, 264637, 18108385 |
| 2077 | 11389877 (4153, 4154) | | | UNCLASSIFIED | 264592 |
| 2078 | 87539364 (4155, 4156) | Novel Protein sim. GBank gij4220590 dbj BAA74579 - (D87908) nuclear protein np95 [Mus musculus] | Contains protein domain (PF00628) - PHD-finger | ubiquitin | 29331825, 265017, 265018, 264288, 265020, 265021, 264634, 56526486 |
| 2079 | 88085916 (4157, 4158) | Novel Protein sim. GBank gij4240255 dbj BAA74906.1 - (AB020690) KIAA0883 protein [Homo sapiens] | Contains protein domain (PF00098) - Zinc finger, CCHC class | UNCLASSIFIED | 264768, 22278997, 265021, 264690, 264259, 264692, 29331822, 264693, 29331824, 29331828, 264508, 264509, 264908, 264907, 264628, 20281089, 264909, 265007, 265009, 264632, 264636, 264591, 264592, 264639, 264758, 264759, 33108954, 264604, 265018, 265019, 22279002, 264563, 264584, 264448, 264684, 264567, 264685 |
| 2080 | 94136689 (4159, 4160) | Novel Protein sim. GBank gij2408021 emb CAB16219.1 - (Z99162) putative vacuolar protein [Schizosaccharomyces pombe] | Contains protein domain (PF01363) - FYVE zinc finger | UNCLASSIFIED | 56182575, 264082, 29331824, 29331826, 29331830, 265017, 265018, 265020, 83373044 |
| 2081 | 94847186 (4161, 4162) | Novel Protein sim. GBank gij5524734 gb AAD44360.1 AF166350 ST7 protein [Homo sapiens] | Contains protein domain (PF00431) - CUB domain | eph | 22278996, 22278997, 22278999, 60432049, 29331822, 29331824, 33657402, 85658542, 265011, 265018, 265019, 21906767, 21906768, 264693, 18108385, 22279000, 22279002 |
| 2082 | 87628629 (4163, 4164) | Novel Protein sim. GBank gij3880558 emb CAA94234 - (Z70271) predicted using GeneFinder; similar to collagen; cDNA EST yk308e7.3 comes from this gene; cDNA EST yk308e7.5 comes from this gene; cDNA EST yk385a8.3 comes from this gene; cDNA EST yk385a8.5 comes from this gene [Caeno...] | | collagen | 264907, 265019 |
| 2083 | 94141000 (4165, 4166) | Novel Protein sim. GBank gij2352427 (AF004161) - peroxisomal Ca-dependent solute carrier [Oryctolagus cuniculus] | Contains protein domain (PF00153) - Mitochondrial carrier proteins | transport | 56182575, 22278996, 22278997, 264259, 29331822, 29331825, 264509, 264112, 265009, 264593, 21906754, 265018, 265019, 264448, 264288, 264685, 21906768, 265020, 265022, 264691, 18108370, 65274791, 264631, 264555, 264563 |
| 2084 | 95199298 (4167, 4168) | Novel Protein sim. GBank gij728836 sp P39193 ALU6_HUMAN - IIII ALU SUBFAMILY SP WARNING ENTRY IIII | | UNCLASSIFIED | 265018, 264763, 264683, 264691 |

| | | | | | |
|------|-----------------------|--|----------|--------------|--|
| 2085 | 94989476 (4169, 4170) | Novel Protein sim. GBank gi 1655699 emb CAA69032 - (Y07752) pherophorin-S [Volvox carten] | | UNCLASSIFIED | 56182575, 60432289, 264908, 56182435, 87168474, 264763, 264369, 264686, 264693, 18108370, 56182323 |
| 2086 | 91234404 (4171, 4172) | Novel Protein sim. GBank gi 3875032 emb CAA88936 - (Z49125) similarity to Trichostrongylus colubriformis 11 kd secretory protein (Swiss Prot accession number P21837); cDNA EST EMBL:D33349 comes from this gene; cDNA EST EMBL:D37644 comes from this gene; cDNA EST EMBL:D36149 come... | | UNCLASSIFIED | 35696286, 264259, 35696052, 264906, 264907, 264908, 264909, 264910, 264759, 264604, 264762, 264768, 264769, 35695917, 263978, 35698423, 35695855, 264632, 264634, 264637, 264638, 264639, 56182323, 18108385, 264482, 264486 |
| 2087 | 21436337 (4173, 4174) | | | UNCLASSIFIED | 264489 |
| 2088 | 94111527 (4175, 4176) | Novel Protein sim. GBank gi 3880930 emb CAA16334.1 - (AL021481) similar to Phosphoglucosylase and phosphonannomutase phosphoserine; cDNA EST EMBL:D36188 comes from this gene; cDNA EST EMBL:D70897 comes from this gene; cDNA EST yk373h9.5 comes from this gene; cDNA EST EMBL:T0080... | | | 264488, 22278994, 35696286, 22278996, 29331827, 35696052, 33657402, 21906754, 33109954, 87168474, 265017, 265018, 265019, 264448, 264683, 264369, 264685, 264687, 264689, 21906765, 21906768, 21906767, 21906768, 21906769, 265020, 265021, 265022, 264692, 33657023, 33657109, 33657182, 27486261, 27486262, 33657349, 27486265, 35696423, 35695855, 83373044, 87168518, 22279000, 264587 |
| 2089 | 95422801 (4177, 4178) | Novel Protein sim. GBank gi 4758118 ref NP_004623.1 pDAP3 - Death associated protein 3 | cadherin | | 18108392, 264488, 52644507, 18108394, 18108397, 52648842, 18108398, 56182575, 22278994, 22278995, 35696286, 22278996, 56994075, 22278997, 22278998, 22278999, 264091, 264092, 264093, 264094, 60432049, 264259, 29331822, 20281099, 29331824, 29331825, 29331826, 29331827, 29331828, 35696052, 33658970, 29146498, 29146499, 264102, 264106, 264107, 264109, 264508, 264905, 264509, 264906, 264907, 264908, 66712502, 264828, 52644045, 264909, 56182435, 264110, 264112, 264510, 264511, 265006, 264512, 265007, 265008, 264910, 265009, 60170831, 264592, 264593, 60433356, 33657402, 80433438, 264595, 55812038, 264758, 21906754, 33657084, 55811386, 52644296, 265010, 265011, 87168559, 265017, 265018, 265019, 264760, 264761, 55811150, 264762, 18108351, 264682, 264448, 264763, 264764, 264683, 264369, 18108354, 264288, 264685, 264766, 264686, 264687, 264768, 52644229, 264688, 18108358, 56181562, 264769, 18108359, 264689, 21906765, 21906766, 21906767, 21906768, 29148627, 21906769, 55811957, 29148629, 29148784, 35695917, 265020, 265021, 265022, 60170615, 264590 |

| | | | | | |
|------|-----------------------|--|---|-------------------|--|
| 2090 | 88222470 (4179, 4180) | | | | 22278995, 22278998, 22278999, 264259, 29331826, 35696052, 264910, 33657402, 60433438, 33109954, 87168474, 87168559, 265018, 265019, 264681, 264684, 264686, 264687, 264688, 264689, 21908765, 21908766, 21908767, 21908769, 35695917, 265022, 60170615, 33657023, 35696423, 35695855, 264952, 18108387, 22279000, 263984, 264905, 264908, 264511, 264512, 265008, 264910, 55811386, 264288, 264768, 56181562, 21906765, 21906768, 21906769, 265022, 264628, 264563, 264567 |
| 2091 | 95309161 (4181, 4182) | Novel Protein sim. GBank gi 4580997 gb AAD24571.1 AF12108 - (AF121081) cAMP inducible 2 protein [Mus musculus] | UNCLASSIFIED | homeobox | 22278997, 22278999, 66712502, 87168559, 264683, 265021, 264486 |
| 2092 | 88223605 (4183, 4184) | | | homeobox | 22278997, 22278999, 66712502, 87168559, 264683, 265021, 264486 |
| 2093 | 87406073 (4185, 4186) | Novel Protein sim. GBank gi 2352427 (AF004161) - peroxisomal Ca-dependent solute carrier [Cryptolegus cuniculus] | Contains protein domain (PF00153) - Mitochondrial carrier proteins | transport | 264905, 264906, 264907, 264908, 264510, 265008, 265007, 265009, 264910, 264596, 21906754, 87168474, 265011, 264603, 265018, 265019, 264760, 264766, 264768, 264769, 21906767, 21906768, 21906769, 265021, 264690, 33657023, 264693, 264628, 264634, 264636, 264637, 264557, 56182323, 264564 |
| 2094 | 91230928 (4187, 4188) | Novel Protein sim. GBank gi 4929551 gb AAD34036.1 AF15179 - (AF151799) CGI-40 protein [Homo sapiens] | | MHC | 35696286, 265017, 265018, 265019, 18108388 |
| 2095 | 95351528 (4189, 4190) | Novel Protein sim. GBank gi 1363238 pir JA57284 - spermatid perinuclear RNA-binding protein Spnr - mouse | Contains protein domain (PF000035) - Double-stranded RNA binding motif | dna_ma_bind | 35696286, 52644045, 265006, 265007, 265008, 87168559, 18108351, 21906769, 29148784, 265020, 33657023, 27486262, 18108374, 18108388 |
| 2096 | 94119760 (4191, 4192) | Novel Protein sim. GBank gi 3834423 (AF070689) - cytoplasmic dynein intermediate chain isoform DIC1a [Drosophila melanogaster] | Contains protein domain (PF00400) - WD domain, G-beta repeat | ATPase_associated | 264488, 264489, 65274572, 56182575, 22278996, 22278997, 22278999, 264259, 60433289, 29331826, 35696052, 264107, 264508, 264509, 264905, 264906, 264907, 264908, 52644045, 264909, 264510, 264511, 264512, 265008, 264910, 265009, 264592, 60433356, 60433438, 264758, 264596, 55812038, 21906754, 264601, 264602, 264605, 264762, 264681, 18108351, 264764, 264683, 264288, 264687, 264768, 264769, 264689, 21906765, 21906766, 21906767, 35695917, 265020, 265022, 52644150, 264691, 264692, 33657023, 264693, 27486261, 35695763, 264628, 264629, 35696423, 35695855, 264631, 264632, 264634, 264635, 264555, 264637, 265981, 264638, 264639, 264563, 264483, 264565, 264566, 264486, 264567 |

| | | | | | |
|------|-----------------------|--|--|------------------|---|
| 2097 | 95322772 (4193, 4194) | Novel Protein sim. GBank gi15174501[ref]NP_006051.1 pLYF1 - zinc finger protein, subfamily 1A, 1 (Ikaros) | Contains protein domain (PF00096) - Zinc finger, C2H2 type | transcriptfactor | 65274572, 264511, 265010, 264600, 265017, 264448, 264288, 265021, 60170815, 264892, 33657109, 18108370, 264636, 264483 |
| 2098 | 87780340 (4195, 4196) | Novel Protein sim. GBank gi1758208[ref]NP_004081.1 pDUSP - dual specificity phosphatase 3 (vaccinia virus phosphatase Vh1-related) | Contains protein domain (PF00782) - Dual specificity phosphatase, catalytic domain | phosphatase | 56994075, 264259, 264288, 265020, 264563 |
| 2099 | 95412827 (4197, 4198) | Novel Protein sim. GBank gi12695659 (AF026954) - pyruvate dehydrogenase phosphatase regulatory subunit precursor, PDP1 [Bos taurus] | | phosphatase | 65274572, 264905, 65274444, 264691, |
| 2100 | 95332656 (4199, 4200) | Novel Protein sim. GBank gi13881189[emb]CAB16514 - (Z99281) similar to ADP-ribosylation factor; cDNA EST EMBL:C08179 comes from this gene; cDNA EST EMBL:C08337 comes from this gene; cDNA EST EMBL:C08829 comes from this gene; cDNA EST yk291b4.5 comes from this gene; cDNA EST yk4... | Contains protein domain (PF00025) - ADP-ribosylation factor family | nucl_recpt | 56182575, 22278995, 22278996, 22278997, 22278998, 60432049, 264259, 26331822, 29331824, 29331825, 29331827, 29331828, 29146488, 264909, 265008, 265009, 264910, 264591, 60432229, 60433356, 33657402, 264758, 21906754, 85658542, 87168474, 265017, 265018, 265019, 264681, 18108351, 264762, 264448, 264369, 264288, 18108355, 264686, 21908765, 21908767, 21908768, 21908769, 265020, 265021, 33657023, 18108374, 35896423, 264558, 83373044, 87168518, 60432113, 22279000, 22279002 |
| 2101 | 87762604 (4201, 4202) | Novel Protein sim. GBank gi14589468[dbj]BAA76761.1 - (AB012808) mBOCT [Mus musculus] | | UNCLASSIFIED | 264091, 29331824, 264105, 265007, 265010, 18108380 |
| 2102 | 87770461 (4203, 4204) | Novel Protein sim. GBank gi13874149[emb]CAA97423.1 - (Z73103) predicted using GeneFinder [Caenorhabditis elegans] | | UNCLASSIFIED | 264488, 264489, 35696286, 264259, 35696052, 264508, 264905, 264907, 264908, 264909, 264511, 264512, 264591, 264593, 60433356, 264758, 264601, 264605, 264760, 18108351, 264448, 264764, 264288, 264767, 264788, 21908769, 35695917, 18108374, 264634, 264555, 264559, 264563, 264482, 264486 |
| 2103 | 95413576 (4205, 4206) | Novel Protein sim. GBank gi14240159[dbj]BAA74858.1 - (AB020842) KIAA0835 protein [Homo sapiens] | Contains protein domain (PF01530) - Zinc finger, C2HC type | transcriptfactor | 65274572, 56994075, 22278999, 264259, 29331824, 29331825, 35696052, 29331828, 66712502, 265009, 60170831, 264595, 33109954, 85658542, 87168559, 265017, 265019, 264448, 21906765, 21906768, 265022, 33657023, 27486262, 33657349, 35695763, 60431528, 18108374, 55811576, 56182323, 18108387, 87168518, 60432113, 264584 |
| 2104 | 85776161 (4207, 4208) | | | UNCLASSIFIED | 264592, 264604, 22279000 |

| | | | | | |
|------|-----------------------|--|--|--------------|--|
| 2105 | 94848080 (4209, 4210) | Novel Protein sim. GBank gi 1707032 (U80445) - coded for by C. elegans cDNA yk13g5.3; coded for by C. elegans cDNA yk21g6.3; coded for by C. elegans cDNA CEMSE18F; coded for by C. elegans cDNA yk126b1.3; coded for by C. elegans cDNA yk65h8.3; coded for by C. elegans cDNA yk65h8.... | | UNCLASSIFIED | 264488, 56182575, 22278994, 56894075, 22278996, 22278997, 22278998, 22278999, 60432049, 264259, 29331822, 29331824, 29331826, 60432289, 29331827, 29331828, 264508, 264905, 264509, 264907, 29331830, 52644045, 264510, 264511, 265007, 264512, 265009, 60170831, 60432228, 33657402, 60433356, 264595, 60433438, 264758, 33657084, 87168474, 265010, 87168559, 265017, 265018, 265019, 264762, 18108351, 264684, 18108354, 264288, 264686, 52644229, 18108359, 21906765, 21906766, 21906767, 21906768, 21906769, 35695917, 265020, 265022, 60170815, 52644150, 264682, 33657023, 264693, 33657109, 60431528, 18108374, 65274791, 35695855, 264635, 60170394, 264639, 264558, 18108385, 18108387, 56526486, 87168518, 60432113, 264564, 264566, 264567, 265006, 265019, 264906, 264639 |
| 2106 | 83365475 (4211, 4212) | Novel Protein sim. GBank gi 3881524[emb]CAA93883] - (Z70038) ZK1067.4 [Caenorhabditis elegans] | | | |
| 2107 | 79822662 (4213, 4214) | Novel Protein sim. GBank gi 3176689 (AC003671) - Contains similarity to ubiquitin carboxyl-terminal hydrolase 14 gb Z35927 from S. cerevisiae. [Arabidopsis thaliana] | | UNCLASSIFIED | 264905, 264906, 264907, 264908, 264909, 264758, 265011, 264600, 264601, 264764, 264766, 264767, 264768, 264769, 264693, 264629, 35695855, 264632, 264634, 264635, 264638, 264639, 83373044, 264488, 18108348, 264769, 18108370, 18108374, 264555, 264556, 264557, 264558, 264564 |
| 2108 | 84233976 (4215, 4216) | Novel Protein sim. GBank gi 481043[pir]S37671 - bat2 protein - human | | UNCLASSIFIED | |
| 2109 | 80478719 (4217, 4218) | Novel Protein sim. GBank gi 2143639[pir]I56542 - calmodulin-binding protein - rat | | UNCLASSIFIED | |
| 2110 | 87729075 (4219, 4220) | Novel Protein sim. GBank gi 4426629[gb]AAD20459] - (AF100980) protocadherin [Rattus norvegicus] | Contains protein domain (PF00069) - struct | UNCLASSIFIED | 264766, 35695917, 264630, 264567, 264486, 264508, 264906, 264591, 264682, 22279002 |
| 2111 | 87818419 (4221, 4222) | Novel Protein sim. GBank gi 4426629[gb]AAD20459] - (AF100980) protocadherin [Rattus norvegicus] | Eukaryotic protein domain (PF00028) - cadherin | | 265006 |
| 2112 | 87283783 (4223, 4224) | Novel Protein sim. GBank gi 3327184[dbj]BAA31660] - (AB014585) KIAA0685 protein [Homo sapiens] | Cadherin domain | | 66714117, 29331826, 29331827, 60433438, 55812038, 265017, 265019, 264689, 21906769, 55811957, 265020, 265021, 33657109, 60170394, 264558 |
| 2113 | 78841388 (4225, 4226) | Novel Protein sim. GBank gi 3327184[dbj]BAA31660] - (AB014585) KIAA0685 protein [Homo sapiens] | | | 85274572, 264689, 264691, 264692, 60432113 |
| 2114 | 87889342 (4227, 4228) | Novel Protein sim. GBank gi 3327184[dbj]BAA31660] - (AB014585) KIAA0685 protein [Homo sapiens] | | UNCLASSIFIED | |
| 2115 | 90993785 (4229, 4230) | Novel Protein sim. GBank gi 4757890[ref]NP_004328.1[pc8OR - chromosome 8 open reading frame 1 | | | |

| | | | | | |
|------|-----------------------|--|--|--------------|--|
| 2116 | 88259387 (4231, 4232) | Novel Protein sim. GBank gii2246532 (U83872) - ORF 73, contains large complex repeat CR 73 (Kaposi's sarcoma-associated herpesvirus) | | struct | 35696286, 22278999, 56182181, 29331824, 29331825, 29331827, 35696052, 284907, 56182435, 265008, 264591, 55812038, 55811386, 87168559, 264288, 264389, 21906769, 29148629, 33657023, 35695763, 55811576, 35698423, 18108385 |
| 2117 | 87788904 (4233, 4234) | Novel Protein sim. GBank gii2330021 (AF019250) - kinesin-related protein; KRP; Costal2 [Drosophila melanogaster] | | struct | 29331824, 264511, 265009, 33109954, 265017, 265018, 264288, 264689, 265020, 264692, 56526486, 264482 |
| 2118 | 87078894 (4235, 4236) | Novel Protein sim. GBank gii1079307 (p1j B56573 - nuclear pore complex glycoprotein p62 - African clawed frog) | | glycoprotein | 264259, 264905, 264907, 264908, 264510, 264511, 265009, 264910, 265010, 264602, 264288, 264768, 264693, 263987, 263972, 264638, 264559 |
| 2119 | 86999317 (4237, 4238) | Novel Protein sim. GBank gii4321407 (gb AAD15748) - (AF047690) ATP-binding cassette protein M-ABC1 [Homo sapiens] | | UNCLASSIFIED | 264693, 18108385 |
| 2120 | 87789395 (4239, 4240) | Novel Protein sim. GBank gii4885527 (ref NP_005480.1 pNSP3 - novel SH2-containing protein 3) | Contains protein domain (PF00017) - eph Src homology domain 2 | eph | 264091, 264259, 29331826, 29331828, 265017, 264604, 264288, 264685, 265020, 264691, 18108370, 55810764, 264555, 264638, 60432113 |
| 2121 | 80021375 (4241, 4242) | Novel Protein sim. GBank gii4757728 (ref NP_004886.1 pAGTA - angiotensin/vasopressin receptor AII/AVP-like) | | UNCLASSIFIED | 264601, 264766, 263978 |
| 2122 | 81230931 (4243, 4244) | Novel Protein sim. GBank gii4929551 (gb AAD34036.1 AF15179 - (AF151799) CGI-40 protein [Homo sapiens]) | | | 18108394, 56182575, 22278997, 29331822, 29331824, 29331825, 29331826, 29331828, 264907, 56182435, 265007, 264910, 265010, 265018, 264686, 265020, 55811576, 264555, 264637, 18108382, 83373044, 18108383, 18108384, 56526486, 264565, 264567 |
| 2123 | 86787898 (4245, 4246) | Novel Protein sim. GBank gii2244551 (dbj BAA20764) - (AB002303) KIAA0305 [Homo sapiens] | Contains protein domain (PF01363) - FYVE zinc finger | struct | 18108386, 264757, 265011, 18108351, 264691, 264634, 18108385 |
| 2124 | 83005951 (4247, 4248) | Novel Protein sim. GBank gii5689455 (dbj BAA83011.1) - (AB028982) KIAA1059 protein [Homo sapiens] | Contains protein domain (PF00801) - PKD domain | transport | 29331822, 264906, 264907, 264591, 264639, 264563 |
| 2125 | 95354041 (4249, 4250) | Novel Protein sim. GBank gii728831 (sp P39188 ALU1_HUMAN - HUMAN - III ALU SUBFAMILY J WARNING ENTRY III) | | UNCLASSIFIED | 264259, 264509, 264907, 264511, 85658542, 264763, 21806765, 35695917, 264636, 264486 |
| 2126 | 95084231 (4251, 4252) | Novel Protein sim. GBank gii4539264 (emb CAB39853.1) - (AL049495) conserved hypothetical protein [Schizosaccharomyces pombe] | | UNCLASSIFIED | 264488, 264489, 29331827, 35696052, 264905, 264509, 264908, 264909, 264510, 265009, 264591, 264592, 264593, 33657402, 264594, 264595, 264596, 264758, 264601, 264603, 265018, 264604, 264605, 264760, 264681, 264762, 264683, 264764, 264684, 264288, 264685, 264689, 60170615, 33657023, 33657109, 55810764, 264635, 264636, 264637, 264638, 264639, 83373044, 264564, 264566 |

| | | | | | |
|------|-----------------------|--|--|-------------------|--|
| 2127 | 81118652 (4253, 4254) | Novel Protein sim. GBank gij4868435[gb AAD31315.1 AF14323] - (AF143236) apoptosis related protein APR-2 [Homo sapiens] | | | 35696286, 28331826, 35696052, 264508, 264509, 264905, 264908, 264907, 264908, 264909, 264510, 265006, 264511, 264512, 265007, 265009, 264910, 264758, 265011, 264600, 264601, 264604, 264762, 264763, 264766, 264687, 264768, 264769, 264689, 35695917, 264690, 264691, 264692, 264693, 264629, 18108374, 35695855, 264632, 264634, 264635, 264636, 264637, 264638, 264639, 18108385, 264563, 264564, 264567, 56182575, 35696286, 56182181, 29331824, 60432289, 35696052, 264905, 264907, 66712502, 264908, 264909, 264510, 264512, 265009, 264910, 264591, 55812038, 265018, 264764, 264288, 264369, 264687, 264768, 55811957, 264692, 18108388, 264628, 264632, 264634, 264635, 264637, 56182323, 264639, 18108384, 18108388, 264563, 264567 |
| 2128 | 87414262 (4255, 4256) | | | | 56714117, 264628, 264595, 55812038, 55811150, 55811957, 264693, 18108374, 263978, 65274791, 18108381, 83373044, 22278000 |
| 2129 | 95102089 (4257, 4258) | | UNCLASSIFIED | | 263981 |
| 2130 | 95417144 (4259, 4260) | Novel Protein sim. GBank gij2649255 (AE001012) - conserved hypothetical protein [Archaeoglobus fulgidus] | UNCLASSIFIED | | |
| 2131 | 85723065 (4261, 4262) | Novel Protein sim. GBank gij1086886 (U41276) - Similar to potassium channel protein. [Caenorhabditis elegans] | Contains protein domain (PF00805) - Pentapeptide repeats (8 copies) | potassium_channel | 35696052, 264909, 264768, 35695917 |
| 2132 | 95361096 (4263, 4264) | Novel Protein sim. GBank gij5699373[dbj BA82973.1 - (AB028944) KIAA1021 protein [Homo sapiens] | Contains protein domain (PF00122) - E1-E2 ATPase | ATPase_associated | 264488, 22278999, 264259, 29331827, 29331828, 35696052, 264509, 264905, 264906, 264907, 264908, 264909, 264510, 265006, 264511, 265007, 264910, 264591, 264592, 264595, 264758, 21906754, 33109954, 87168474, 265011, 264600, 264601, 264605, 265019, 264760, 18108351, 264681, 264762, 264764, 264288, 264684, 264766, 264686, 264687, 264768, 264769, 264688, 21908769, 264690, 52644150, 264691, 264693, 18108370, 264628, 264629, 18108372, 18108374, 35696423, 35695955, 264631, 264634, 264635, 264636, 264555, 264637, 18108380, 264639, 264558, 56182323, 56526486, 264564, 264565, 264566, 264567 |

| | | | | | |
|------|-----------------------|---|--|-------------------|--|
| 2133 | 95351539 (4265, 4266) | Novel Protein sim. GBank gij4220489 (AC006069) - hypothetical protein [Arabidopsis thaliana] | | UNCLASSIFIED | 60424179, 52646365, 52646842, 56894075, 35696286, 22278997, 22278998, 60432049, 56182181, 66714117, 60424269, 29331826, 29331828, 35696052, 264905, 264906, 264907, 66712502, 29331830, 58182435, 265006, 264512, 265008, 60431735, 60433356, 33657402, 55812038, 33109954, 21906754, 55811386, 265010, 264603, 265017, 265018, 265019, 55811150, 18108351, 264682, 264369, 264288, 52644229, 56181562, 21906785, 21908766, 21906767, 21906768, 21906769, 55811957, 35695917, 265020, 265021, 60170615, 33657023, 33657109, 60431528, 18108374, 35696423, 65274791, 35695855, 264634, 60431850, 18108380, 56182323, 83373044, 18108385, 18108387, 60432113, 22279002, 264563, 264566 |
| 2134 | 95412697 (4267, 4268) | Novel Protein sim. GBank gij3875351[emb](CAB09415) - (Z96047) DY3.6 [Caenorhabditis elegans] | | | 56181686, 35696286, 21906754, 55811386, 265011, 265017, 18108351, 264785, 264766, 264688, 21906768, 35695917, 265020, 33657023, 264628, 35695855, 264632, 264555, 264556, 264557, 264558, 18108382, 22279002 |
| 2135 | 88079813 (4269, 4270) | Novel Protein sim. GBank gij5689559[dbj](BAA83063.1) - (AB029034) KIAA1111 protein [Homo sapiens] | Contains protein domain (PF00628) - PHD-finger | UNCLASSIFIED | 22278999, 29331828, 35696052, 264906, 264908, 264910, 265009, 264591, 264758, 52646317, 265011, 87188559, 264601, 18108351, 284448, 264683, 264684, 264689, 18108359, 264691, 33657023, 264692, 35695763, 264629, 35695855, 264631, 264635, 264636, 264637, 56182323, 264639, 22279002, 264584 |
| 2136 | 84346479 (4271, 4272) | Novel Protein sim. GBank gij2662167[dbj](BAA23715) - (AB007903) KIAA0443 [Homo sapiens] | | UNCLASSIFIED | 264639 |
| 2137 | 87637716 (4273, 4274) | Novel Protein sim. GBank gij4884110[emb](CAB43262.1) - (AL050090) hypothetical protein [Homo sapiens] | | UNCLASSIFIED | 264569, 264909, 33109954, 264763, 21906768, 60170394, 18108385, 264563 |
| 2138 | 87395446 (4275, 4276) | Novel Protein sim. GBank gij5174779[gb](AAD40686.1) - (U87804) 50 kDa protein [Caulobacter crescentus] | | ATPase_associated | 264259, 29331828, 35696052, 264909, 265006, 265017, 265018, 18108351, 264288, 21906768, 33657023, 33657109, 264628, 18108374, 35695855, 264634, 264555, 264556, 264557, 264558, 264559 |
| 2139 | 94843882 (4277, 4278) | Novel Protein sim. GBank gij3850821[emb](CAA77135) - (Y18350) U2 snRNP auxiliary factor, large subunit [Nicotiana glauca] | | | 264905, 264910, 264591, 55812038, 55811386, 8568542, 264760, 18108351, 18108359, 55811957, 265020, 265021, 33657023, 18108364, 55811576, 83373044, 18108385, 56526486, 264482 |

| | | | | |
|------|-----------------------|--|--------------|--|
| 2140 | 87645655 (4279, 4280) | Novel Protein sim. GBank gij4417293jgb/AAD20418] - (AC007018) unknown protein [Arabidopsis thaliana] | UNCLASSIFIED | 264488, 264259, 29331824, 264104, 264109, 264509, 265006, 264759, 265018, 264448, 264288, 21908768, 55811957, 265021, 33657023, 27488265, 35696423, 264636, 264556, 264557, 264559, 264566 |
| 2141 | 78623988 (4281, 4282) | | UNCLASSIFIED | 265020, 264693 |
| 2142 | 80041222 (4283, 4284) | | UNCLASSIFIED | 263978 |
| 2143 | 94140051 (4285, 4286) | Novel Protein sim. GBank gij2135766jpir/S53362 - mucin 5AC (clone JER47) - human (fragment) | UNCLASSIFIED | 22278997, 29331827, 264907, 265020, 60432113 |
| 2144 | 94320114 (4287, 4288) | Novel Protein sim. GBank gij2076483 (U43200) - antifreeze glycoprotein AFGP polyprotein precursor [Boreogadus saida] | UNCLASSIFIED | 65274572, 264259, 29331824, 29331827, 264906, 264908, 264591, 265011, 87168559, 264600, 265019, 264288, 264768, 21908765, 21908767, 55811578, 35696423, 65274791, 22279002 |
| 2145 | 20564305 (4289, 4290) | | UNCLASSIFIED | 263978 |
| 2146 | 87010515 (4291, 4292) | Novel Protein sim. GBank gij1255871 (U53341) - short region of weak similarity to bovine membrane receptor p63 (PIR:S28503) [Caenorhabditis elegans] | UNCLASSIFIED | 264908, 60433356, 264686 |
| 2147 | 80432911 (4293, 4294) | Novel Protein sim. GBank gij3080398jemb/CAA18718.1] - (AL022603) putative protein [Arabidopsis thaliana] | UNCLASSIFIED | 264907, 264768, 264769, 18108385 |
| 2148 | 80048811 (4295, 4296) | Novel Protein sim. GBank gij728837jpiP39194ALU7_HUMAN - III ALU SUBFAMILY SQ WARNING ENTRY IIII | UNCLASSIFIED | 264593 |
| 2149 | 87362022 (4297, 4298) | Novel Protein sim. GBank gij119863jpiP20693jFCE2_MOUSE - LOW AFFINITY IMMUNOGLOBULIN EPSILON FC RECEPTOR (LYMPHOCYTE IGE RECEPTOR) (FC-EPSILON-RII) (CD23) | glycoprotein | 29331824, 29331826, 35696052, 264758, 87168474, 265018, 52644150, 33657109 |
| 2150 | 94140059 (4299, 4300) | Novel Protein sim. GBank gij5420387jemb/CAB46678.1] - (AJ243459) proteophosphoglycan [Leishmania major] | UNCLASSIFIED | 22278998, 29331822, 29331824, 29331828, 264764, 264769, 21908766, 264486 |
| 2151 | 95353241 (4301, 4302) | Novel Protein sim. GBank gij5689407jdbj/BAA82987.1] - (AB028958) KIAA1035 protein [Homo sapiens] | | 22278996, 56994075, 22278999, 60432049, 264259, 29331822, 29331824, 29331826, 35696052, 29331828, 264508, 264511, 60433356, 264758, 264596, 33109954, 60174639, 265010, 265011, 87168559, 265017, 265018, 265019, 264448, 264288, 264689, 21908765, 21908768, 21908768, 265020, 60170615, 33657109, 33657182, 33657349, 18108370, 264635, 264557, 60170394, 18108385, 87168518, 22279000, 18108397, 18108398, 265007, 264591, 265011, 18108351, 18108368, 18108374, 18108388 |
| 2152 | 79321640 (4303, 4304) | Novel Protein sim. GBank gij3452473 (AF084205) - serine/threonine protein kinase TAO1 [Rattus norvegicus] | kinase | |

| | | | | | |
|------|-----------------------|---|---|--------------|--|
| 2153 | 88313371 (4305, 4306) | Novel Protein sim. GBank gi14758704[refNP_004216.1]pMASL - MFH-amplified sequences with leucine-rich tandem repeats 1 | Contains protein domain (PF00560) - Leucine Rich Repeat | glycoprotein | 264488, 263994, 52646842, 22278996, 22278998, 22278999, 264259, 29331822, 35698052, 264508, 264509, 264905, 264906, 264907, 264908, 264909, 56182435, 264510, 264511, 264512, 264758, 87168474, 87168559, 265017, 265019, 264760, 264288, 264369, 264766, 264687, 264769, 52644229, 21906768, 21906768, 35695917, 33657023, 33657109, 35695855, 264631, 264632, 264635, 264636, 264639, 18108385, 264483, 264554, 264488 |
| 2154 | 87408034 (4307, 4308) | Novel Protein sim. GBank gi225150[prf1209265U - chorion protein B11 [Bombyx mori] | | UNCLASSIFIED | 56994075, 264094, 265009, 265019, 264288, 21908767, 35695917 |
| 2155 | 87424072 (4309, 4310) | | | UNCLASSIFIED | 18108392, 18108398, 22278996, 264259, 29331824, 265008, 265010, 265011, 265017, 265019, 264288, 264686, 265020, 264693, 264628, 56182323 |
| 2156 | 84295205 (4311, 4312) | Novel Protein sim. GBank gi3970966 (AC004974) - spa-1- (like, similar to AF026504 (PID:g2555183) [Homo sapiens]) | | | 265007, 264694 |
| 2157 | 87316344 (4313, 4314) | Novel Protein sim. GBank gi1076211[pir1S50755 - hypothetical protein VSP-3 - Chlamydomonas reinhardtii] | | UNCLASSIFIED | 264591 |
| 2158 | 86444218 (4315, 4316) | Novel Protein sim. GBank gi4650844[dbjBAA77027.1] - (AB026190) Kelch motif containing protein [Homo sapiens] | | UNCLASSIFIED | 264596 |
| 2159 | 80083729 (4317, 4318) | Novel Protein sim. GBank gi2879925[dbjBAA24826] - (AB007897) KIAA0437 [Homo sapiens] | Contains protein domain (PF00651) - BTB/POZ domain | dna_rna_bind | 29331822, 264112, 265009, 264691, 33657023, 264634 |
| 2160 | 16283674 (4319, 4320) | Novel Protein sim. GBank gi2879925[dbjBAA24826] - (AB007897) KIAA0437 [Homo sapiens] | | | 264634 |
| 2161 | 87739131 (4321, 4322) | Novel Protein sim. GBank gi1504006[dbjBAA13202] - (D88968) simliarto human ZFY protein. [Homo sapiens] | | UNCLASSIFIED | 265008 |
| 2162 | 84318526 (4323, 4324) | Novel Protein sim. GBank gi1504006[dbjBAA13202] - (D88968) simliarto human ZFY protein. [Homo sapiens] | | UNCLASSIFIED | 65274572, 264508, 264805, 264906, 264907, 264808, 52644045, 264909, 265007, 264910, 264591, 264592, 264593, 55812038, 264596, 264758, 265011, 264600, 264762, 264763, 264683, 264764, 264288, 264766, 264686, 264768, 264769, 264689, 265020, 264691, 264628, 264629, 263978, 264632, 264634, 264557, 264638, 264639, 18108385, 264563, 264566, 264567 |
| 2163 | 85417158 (4325, 4326) | Novel Protein sim. GBank gi3876537[embCAA98270] - (Z73974) cDNA EST yk291f5.3 comes from this gene. cDNA EST yk291f5.5 comes from this gene [Caenorhabditis elegans] | | UNCLASSIFIED | 56182575, 22278996, 264093, 264683, 33657023, 65274620, 60432113 |
| 2164 | 80559456 (4327, 4328) | | Contains protein domain (PF01006) - Hepatitis C virus non-structural protein NS4a | collagen | 264503, 264637, 264565 |

| | | | | | |
|------|-----------------------|--|---|--------------|--|
| 2165 | 84329169 (4329, 4330) | Novel Protein sim. GBank gi 1086794 (U41107) - No definition line found [Caenorhabditis elegans] | | UNCLASSIFIED | 56994075, 22278996, 22278997, 22278998, 264259, 29331822, 29331824, 29331825, 29331826, 29331827, 29331828, 264906, 29331830, 56182435, 265009, 21906754, 33657084, 265011, 265019, 264448, 264288, 264369, 21906765, 21906768, 21906769, 265020, 265021, 264691, 264692, 33657023, 65274620, 35695855, 264556, 60170394, 83373044, 60432113, 22279002, 264567 |
| 2166 | 87618934 (4331, 4332) | Novel Protein sim. GBank gi 2706522 emb CAA75816 - (Y15885) ubiquitin activating enzyme [Drosophila melanogaster] | ubiquitin | | 52645158, 22278994, 22278998, 66714117, 29331828, 52644045, 265018, 265019, 264369, 21906765, 21906767, 21906768, 21906769, 265021, 265022, 264693, 27486262, 35695763, 18108376, 56526486, 87168518, 264567 |
| 2167 | 87716864 (4333, 4334) | Novel Protein sim. GBank gi 2224713 dbj BAA20840 - (AB002384) KIAA0386 [Homo sapiens] | | UNCLASSIFIED | 56182575, 35696286, 29331824, 29331826, 29146488, 56182435, 265008, 265009, 264592, 264593, 33657402, 33109954, 265011, 265017, 265018, 18108351, 264369, 21906764, 21906765, 21906768, 29148627, 21906769, 52644150, 33657109, 35696423, 18108381, 18108384, 18108385, 60432113, 264567 |
| 2168 | 86999334 (4335, 4336) | Novel Protein sim. GBank gi 4321407 gb AAD15748 - (AF047690) ATP-binding cassette protein M-ABC1 [Homo sapiens] | Contains protein domain (PF00664) - ABC transporter transmembrane region. | transport | 66714117, 29331827, 264907, 264511, 264591, 265018, 264764, 264683, 264766, 264768, 264568 |
| 2169 | 87886937 (4337, 4338) | Novel Protein sim. GBank gi 5106521 gb AAD39741.1 AF10536 - (AF105365) K-CI cotransporter KCC4 [Homo sapiens] | | UNCLASSIFIED | 264629, 264555, 264559 |
| 2170 | 94141033 (4339, 4340) | | | | 65274572, 56182575, 22278997, 22278998, 264259, 29331825, 264509, 264906, 56182435, 60433438, 55812038, 264596, 55811386, 265019, 264762, 264763, 264448, 264764, 264684, 264288, 264766, 264685, 56181562, 264689, 55811957, 265020, 264535, 264691, 33657109, 60431528, 18108374, 35696423, 55811576, 65274791, 264634, 264639, 264558, 87168518, 60432113, 264564 |
| 2171 | 80194050 (4341, 4342) | | | UNCLASSIFIED | 264369, 285020, 264558 |
| 2172 | 85452460 (4343, 4344) | | | UNCLASSIFIED | 264259, 264558 |
| 2173 | 87036740 (4345, 4346) | Novel Protein sim. GBank gi 4309681 gb AAD15478 - (AC008930) R33423.1 [Homo sapiens] | | UNCLASSIFIED | 264369 |
| 2174 | 95003288 (4347, 4348) | Novel Protein sim. GBank gi 2493778 sp Q08456 YQ35_CAEEL - PUTATIVE CUTICLE COLLAGEN C09G5.5 | | | 264908, 35695855, 264555, 264557 |

| | | | | | |
|------|-----------------------|--|---|--------------|--|
| 2175 | 94325850 (4349, 4350) | Novel Protein sim. GBank gi1263287 (U47855) - fibroin-3 (Araneus diadematus) | | UNCLASSIFIED | 264488, 35696286, 20281099, 29331826, 60432289, 35696052, 264109, 264508, 264509, 264905, 264906, 264907, 264908, 264909, 264510, 264511, 265006, 265007, 264512, 265008, 264910, 264591, 264594, 264595, 264596, 264758, 55812038, 265011, 264600, 264603, 264760, 264762, 264448, 264764, 264288, 264766, 264686, 264687, 21906768, 55811957, 35695917, 265020, 265022, 264691, 264692, 33657023, 264693, 264628, 264629, 55811576, 35696423, 65274791, 35695855, 264631, 264632, 264634, 264635, 264636, 264637, 264638, 264639, 264558, 18108385, 60432113, 264563, 264564, 264565, 264566, 264486, 264567 |
| 2176 | 88223392 (4351, 4352) | Novel Protein sim. GBank gi1728837 sp P39184 ALU7_HUMAN - IIII ALU SUBFAMILY SQ WARNING ENTRY IIII | Contains protein domain (PF00805) - Pentapeptide repeats (8 copies) | oncogene | 52644507, 52646842, 22278994, 35696286, 22278996, 22278999, 29331828, 29331827, 35696052, 29331828, 33656970, 29331830, 264910, 33657402, 264758, 52644296, 87168559, 265018, 264688, 21906765, 21906767, 21906769, 35695917, 52644150, 264690, 33657023, 33657109, 52645129, 33657182, 27486261, 27486262, 33657349, 18108376, 18108377, 35695855, 87168518, 60432113, 264404, 22279000, 264486 |
| 2177 | 94128942 (4353, 4354) | Novel Protein sim. GBank gi15454072 ref NP_006416.1 pSLU7 - step II splicing factor SLU7 | | kinase | 18108392, 22278997, 22278999, 264093, 33657402, 265019, 264448, 264766, 264689, 21906767, 21906768, 21906769, 265021, 33657023, 18108370, 18108374, 60432113, 22279002 |
| 2178 | 87601557 (4355, 4356) | Novel Protein sim. GBank gi1473407 (U08215) - NST-1 (Mus musculus) | Contains protein domain (PF00012) - Hsp70 protein | eph | 264488, 22278996, 22278999, 29331824, 29331825, 29331826, 29331827, 52644296, 87168474, 18108370, 35695855, 22279002 |
| 2179 | 87316275 (4357, 4358) | | | UNCLASSIFIED | 60424269, 264760, 264628, 264632 |

| | | | | | |
|------|-----------------------|--|---|-------------------|---|
| 2180 | 95351397 (4359, 4360) | Novel Protein sim. GBank gi 3122317 sp P80648 KMH_B_DICD1 - MYOSIN HEAVY CHAIN KINASE B (MHCK B) | Contains protein domain (PF00400) - WD domain, G-beta repeat | kinase | 52644507, 22278994, 35696286, 22278997, 22278999, 284259, 52645080, 29331822, 29331824, 29331825, 29331826, 29331828, 33656970, 284508, 264509, 264908, 29331830, 264909, 264510, 265008, 265007, 33657402, 55812038, 21906754, 87168474, 87168559, 265017, 265018, 265019, 264763, 264682, 264683, 264684, 264288, 264686, 21906765, 21906766, 21906769, 265020, 265021, 265022, 52644150, 33657023, 33657109, 27486265, 33657349, 18108374, 35696423, 35695855, 263981, 60170394, 18108385, 58528486, 87168518, 60432113, 22279000, 264482, 264566, 264567, 264486 29331827, 264369, 18108376, 264564 |
| 2181 | 85764930 (4361, 4362) | Novel Protein sim. GBank gi 3024689 sp Q15542 T2D4_HUMAN - TRANSCRIPTION INITIATION FACTOR TFIID 100 KD SUBUNIT (TAFII-100) (TAFII100) | | kinase | |
| 2182 | 87637731 (4363, 4364) | Novel Protein sim. GBank gi 5420387 emb CAB46679.1 - (AJ243459) proteophosphoglycan [Leishmania major] | | UNCLASSIFIED | 22278986, 22278997, 22278999, 264259, 29331822, 56182435, 264112, 264764, 264288, 21906767, 21906768, 21906769, 33657109, 18108376, 60170394, 22279000, 22279002 264760 |
| 2183 | 85460649 (4365, 4366) | Novel Protein sim. GBank gi 3873406 gb AAC77482.1 - (U17129) unknown [Rhodococcus erythropolis] | | inf | 29331822, 29331825, 29331826, 56182435, 265011, 264685, 264686, 21906768, 18108370, 264629, 264631, 264636, 264557 |
| 2184 | 87760690 (4367, 4368) | Novel Protein sim. GBank gi 3114713 (AF061346) - Edp1 protein [Mus musculus] | | | 29331824, 264907, 66712502, 264757, 265019, 264288, 264692, 56526486 |
| 2185 | 87826463 (4369, 4370) | Novel Protein sim. GBank gi 5106956 gb AAD39906.1 AF11361 - (AF113615) FH1/FH2 domain-containing protein FHOS [Homo sapiens] | | ATPase associated | 264259, 29331822, 29331824, 29331826, 56182435, 264592, 55812038, 264760, 264766, 55811957, 33657023, 55811576, 56182323, 264563 |
| 2186 | 87739227 (4371, 4372) | Novel Protein sim. GBank gi 2864625 emb CAA16972 - (AL021811) putative protein [Arabidopsis thaliana] | | UNCLASSIFIED | 35696052, 264905, 264908, 264907, 264908, 264510, 264511, 265008, 264910, 264758, 265019, 264762, 264681, 264766, 264769, 35695917, 264692, 35696423, 264831, 264535, 264837, 18108388, 264586, 264486 |
| 2187 | 87388173 (4373, 4374) | | | histone | 18108398, 56994075, 264259, 29331824, 29331825, 66714117, 29331827, 264908, 29331830, 265018, 265020, 265021, 56182323, 264559, 22279000, 22279002 |
| 2188 | 87771708 (4375, 4376) | Novel Protein sim. GBank gi 5107816 gb AAD40129.1 AF14941 - (AF149413) contains similarity to histone deacetylases; Pfam PF00850, Score=13.3, E=5e-10, N=1 [Arabidopsis thaliana] | | | |
| 2189 | 85693573 (4377, 4378) | Novel Protein sim. GBank gi 3452357 (AF075724) - unknown [Legionella pneumophila] | Contains protein domain (PF01596) - O-methyltransferase | | 22278996, 264259, 29331826, 21906754, 264369, 264288, 263967 |

| | | | | | |
|------|-----------------------|---|---|---------------|--|
| 2190 | 87639197 (4379, 4380) | Novel Protein sim. GBank gij132575 sp P29315 RIN1_RAT - RIBONUCLEASE INHIBITOR | | nucleaseinhib | 22278996, 22278999, 29331822, 29331824, 29331826, 265008, 264910, 60170831, 55812038, 52644296, 265010, 265018, 264685, 264688, 56181562, 21906769, 35695917, 265022, 60170394, 22279000 |
| 2191 | 95198928 (4381, 4382) | Novel Protein sim. GBank gij5327002 emb CAB46272.1 - (Y18503) XAP-5-like protein [Homo sapiens] | | | 29331825, 29331826, 29331830, 264510, 264511, 264910, 264593, 264594, 264556, 264559, 264558 |
| 2192 | 11126316 (4383, 4384) | Novel Protein sim. GBank gij462600 sp P34400 M110_CAEEL - MIG-10 PROTEIN | Contains protein domain (PF00169) - PH domain | | |
| 2193 | 94140073 (4385, 4386) | Novel Protein sim. GBank gij5420388 emb CAB46680.1 - (AJ243460) proteophosphoglycan [Leishmania major] | | UNCLASSIFIED | 56181866, 29331825, 29331827, 264508, 264909, 265008, 264592, 60432229, 264288, 264684, 264766, 35695917, 33657023, 60431602, 60431528, 55810764, 55811576, 65274791, 35695855, 60431850, 56182323, 60432113, 264592 |
| 2194 | 21418714 (4387, 4388) | Novel Protein sim. GBank gij2773341 (AF040954) - putative protein phosphatase 1 nuclear targeting subunit [Rattus norvegicus] | | | |
| 2195 | 88083023 (4389, 4390) | Novel Protein sim. GBank gij2832763 emb CAA15685.1 - (AL009191) /prediction=(method:; /prediction=(method:; /match=(desc:; /motif=(desc: [Drosophila melanogaster] | | UNCLASSIFIED | 22278996, 22278999, 35696052, 265006, 21906754, 265017, 35695917, 265021, 265022, 35695855 |
| 2198 | 85091631 (4391, 4392) | Novel Protein sim. GBank gij5262487 emb CAB45699.1 - (AL080076) hypothetical protein [Homo sapiens] | | collagen | 56182575, 35696286, 22278997, 22278999, 264259, 29331822, 66714117, 60432289, 29331827, 35696052, 29331828, 264508, 52844045, 56182435, 264510, 265007, 265008, 265009, 60433438, 55812038, 265010, 265011, 264448, 264288, 264686, 264687, 52644229, 21906765, 21906766, 21906767, 35695917, 265022, 264691, 33657023, 264693, 18108370, 18108376, 35696423, 55811576, 65274791, 35695855, 264636, 56182323, 18108385, 264768, 264769, 21906765, 21906766, 21906767, 29148627, 55811957, 35696286, 265020, 22278998, 265021, 264259, 33657023, 264693, 29331824, 35696052, 29331828, 18108370, 35695855, 264113, 265008, 264910, 60432229, 56182323, 33657402, 264758, 83373044, 21906754, 265018, 265019, 22279002, 264482, 264448, 264565, 264288, 264369 |
| 2197 | 95073813 (4393, 4394) | Novel Protein sim. GBank gij4929567 gb AAD34044.1 AF15180 - (AF151807) CGI-49 protein [Homo sapiens] | | | |
| 2198 | 88060914 (4395, 4396) | Novel Protein sim. GBank gij3548787 (AC005622) - R30953.1 [Homo sapiens] | | UNCLASSIFIED | |

| | | | | | | |
|------|-----------------------|--|---------------|---|--|--|
| 2199 | 86054355 (4397, 4398) | Novel Protein sim. GBank gjl2739372 (AC002505) - hypothetical protein [Arabidopsis thaliana] | | | | 264105, 264110, 264112, 264688, 55811957, 33657023, 264692, 263967, 20281071, 56526486 |
| 2200 | 87405385 (4399, 4400) | Novel Protein sim. GBank gjl3043634 [dbj BAA25481] - (AB011127) KIAA0555 protein [Homo sapiens] | struct | | | 29331824, 264763, 264768 |
| 2201 | 94316872 (4401, 4402) | Novel Protein sim. GBank gjl3913470 [sp O57314 DHXB_ANAPL - PUTATIVE STEROID DEHYDROGENASE SPM2] | dehydrogenase | Contains protein domain (PF00106) - short chain dehydrogenase | | 29331824, 35696052, 264905, 264907, 33657402, 55811386, 265017, 265018, 265019, 264288, 21906788, 35695917, 265020, 265022, 33657023, 33657109, 27486261, 18108370, 35696423, 35695855, 264555, 264556, 83373044, 87168518, 60432113 |
| 2202 | 91672385 (4403, 4404) | Novel Protein sim. GBank gjl5262665 [emb CAB45767.1] - (AL080186) hypothetical protein [Homo sapiens] | UNCLASSIFIED | | | 264489, 264259, 29331824, 60432289, 35696052, 264905, 264909, 264592, 265017, 265018, 265019, 18108351, 264762, 284448, 264369, 264288, 264766, 21906765, 21906768, 264690, 264691, 264692, 33657109, 264634, 264636, 264555, 264639, 264558, 264559, 83373044, 18108385, 264404, 22279002, 264482 |
| 2203 | 87761832 (4405, 4406) | Novel Protein sim. GBank gjl1172845 [sp P46629 RB25_RABIT - RAS-RELATED PROTEIN RAB-25] | glycoprotein | Contains protein domain (PF00071) - Ras family | | 52646365, 56994075, 264259, 29331822, 29331826, 29331827, 29331828, 264910, 265010, 265011, 87168559, 265018, 265019, 264605, 264288, 21906769, 35695917, 33657023, 264692, 33657109, 35695763, 18108376, 264638, 22279000, 264566, 264567 |
| 2204 | 88088671 (4407, 4408) | Novel Protein sim. GBank gjl121035 [sp P29348 GBT3_RAT GUANINE NUCLEOTIDE-BINDING PROTEIN G(T), ALPHA-3 SUBUNIT (GUSTDUCIN ALPHA-3 CHAIN)] | UNCLASSIFIED | Contains protein domain (PF00503) - G-protein alpha subunit | | |
| 2205 | 84147589 (4409, 4410) | Novel Protein sim. GBank gjl4589480 [dbj BAA76768.1] - (AB023141) KIAA0924 protein [Homo sapiens] | dna_rna_bind | Contains protein domain (PF00096) - Zinc finger, C2H2 type | | 18108394, 18108397, 56182575, 60432049, 264259, 29331822, 29331824, 29331825, 29331826, 29331827, 264906, 265007, 265008, 265009, 60432229, 265010, 265011, 265018, 264683, 264288, 264369, 264686, 21906766, 21906768, 21906769, 264690, 264691, 264693, 18108388, 55811576, 65274791, 264634, 18108381, 18108384, 60432113, 22278002, 264563, 264568 |
| 2206 | 20620008 (4411, 4412) | Novel Protein sim. GBank gjl4557753 [ref NP_000372.1 pMID1 - midline 1 protein] | UNCLASSIFIED | Contains protein domain (PF00622) - SPRY domain | | 264591 |
| 2207 | 87787870 (4413, 4414) | | | | | 29331822, 56182181, 29331827, 35696052, 52644045, 265006, 265019, 56181562, 55811957, 265021, 33657023, 35695763, 35695855, 60170394, 60432113, 264568 |
| 2208 | 86100830 (4415, 4416) | Novel Protein sim. GBank gjl3986746 (AF105228) - tuftelin [Bos taurus] | struct | | | 264906, 265019, 18108351, 21906769 |
| 2209 | 87800420 (4417, 4418) | | | | | 264112, 265009, 264691, 18108365, 18108374, 264634, 20281166 |

| | | | | | |
|------|-----------------------|--|---|----------|---|
| 2210 | 57152407 (4418, 4420) | Novel Protein sim. GBank gij728837 sp P39194 ALU7_HUMAN - IIII ALU SUBFAMILY SQ WARNING ENTRY IIII | | kinase | 264603 |
| 2211 | 87341720 (4421, 4422) | Novel Protein sim. GBank gij728837 sp P39194 ALU7_HUMAN - IIII ALU SUBFAMILY SQ WARNING ENTRY IIII | | oncogene | 264685, 264686, 18108365, 22279002, 264482 |
| 2212 | 91223924 (4423, 4424) | Novel Protein sim. GBank gij3776027 emb CAA09214 - (AJ010475) RNA helicase [Arabidopsis thaliana] | Contains protein domain (PF00270) - DEAD/DEAH box helicase | helicase | 22278995, 22278997, 22278999, 264082, 264094, 29331822, 66714117, 29331826, 29331828, 264907, 52644045, 265009, 60170831, 21906754, 87168559, 265017, 265019, 18108351, 264683, 18108354, 264369, 264766, 264687, 52644229, 21906765, 21906766, 21906767, 21906768, 265021, 33657109, 18108370, 18108374, 264638, 56182323, 18108384, 18108387, 87168518, 264565 |
| 2213 | 91219309 (4425, 4426) | Novel Protein sim. GBank gij5420387 emb CAB46679.1 - (AJ243459) proteophosphoglycan [Leishmania major] | | | 56182575, 22278996, 22278997, 35696052, 264905, 66712502, 264908, 264828, 56182435, 264112, 265008, 60431735, 60433438, 21906754, 265010, 265011, 265017, 265018, 265019, 18108351, 264765, 21906765, 21906768, 21906769, 265020, 265021, 264693, 264629, 263974, 263976, 18108379, 55811576, 264556, 264637, 264558, 83373044, 22278002, 264482, 264483 |

| | | | | | |
|------|-----------------------|--|---|---------------|---|
| 2214 | 95361453 (4427, 4428) | Novel Protein sim. GBank gi 4504325 ref NP_000173.1 pHADH - hydroxyacyl- Coenzyme A dehydrogenase/3-ketoacyl-Coenzyme A thiolase/enoyl-Coenzyme A hydratase (trifunctional protein). alpha su | Contains protein domain (PF00725) - 3-hydroxyacyl-CoA dehydrogenase | dehydrogenase | 264488, 52644507, 18108394, 56182575, 22278994, 22278995, 35696286, 56994075, 22278997, 22278998, 22278999, 264490, 60432049, 264259, 52645080, 29331822, 29147620, 29331824, 66714117, 29331825, 60432289, 29331826, 29331827, 35696052, 29331828, 20281100, 264509, 264807, 66712502, 264908, 29331830, 52644045, 56182435, 264510, 265006, 264511, 264512, 265007, 265008, 265009, 60170831, 264593, 60433356, 60433438, 33109954, 33657084, 52644298, 87168474, 265010, 265011, 87168559, 264601, 265017, 265018, 265019, 18108351, 264448, 264682, 264763, 264288, 264687, 52644229, 264689, 21906765, 21906766, 21906767, 21906768, 21906769, 35695917, 265020, 265021, 265022, 264532, 60170815, 264690, 52644150, 264691, 33657023, 264692, 18108364, 33657109, 33657182, 27486262, 27486264, 27486265, 35695763, 18108370, 264629, 60431528, 18108374, 18108376, 55810764, 35696423, 35695855, 264634, 264636, 52644332, 264638, 264558, 60170394, 18108381, 56182323, 83373044, 18108385, 18108387, 18108388, 56526486, 87168518, 60432113, 22279002, 264482, 264564, 264565, 264566, 264909, 265008, 264555, 264558, 87168518 |
| 2215 | 95418208 (4429, 4430) | Novel Protein sim. GBank gi 1947160 (AF000298) - weak similarity to collagens; glycine- and proline-rich [Caenorhabditis elegans] | | | |
| 2216 | 87614046 (4431, 4432) | Novel Protein sim. GBank gi 1572802 (U70854) - similar to Enterococcus faecalis TRAB (GI:388268) [Caenorhabditis elegans] | | UNCLASSIFIED | 264693 |
| 2217 | 80589404 (4433, 4434) | Novel Protein sim. GBank gi 5031707 ref NP_005503.1 pGARP - glycoprotein A repetitions predominant | Contains protein domain (PF00560) - Leucine Rich Repeat | glycoprotein | 264288, 33657109, 264556 |
| 2218 | 85518254 (4435, 4436) | Novel Protein sim. GBank gi 3878636 emb CAA88953 - (Z49128) similar to cAMP-dependant protein kinase; cDNA EST EMBL: T00719 comes from this gene; cDNA EST yk465d8.3 comes from this gene; cDNA EST yk465d8.5 comes from this gene; cDNA EST yk492d4.3 comes from this gene; cDNA EST y... | Contains protein domain (PF00069) - Eukaryotic protein kinase domain | | 35696423, 264563 |
| 2219 | 87614048 (4437, 4438) | Novel Protein sim. GBank gi 1572802 (U70854) - similar to Enterococcus faecalis TRAB (GI:388268) [Caenorhabditis elegans] | Contains protein domain (PF01963) - TraB family | | 264682, 264683, 264688, 264689, 264693, 18108370, 18108376 |

| | | | | | |
|------|-----------------------|--|---|--------------|---|
| 2220 | 85354165 (4439, 4440) | Novel Protein sim. GBank gi 4507281 ref NP_003145.1 pSTAT - statherin | | | 264488, 18108394, 18108395, 35696286, 264259, 264097, 60432289, 264509, 264905, 264906, 264907, 29331830, 264908, 264909, 264510, 264511, 265007, 264512, 264910, 265009, 264593, 264594, 60433356, 264595, 55812038, 264758, 85658542, 265010, 264601, 264603, 265019, 264605, 264760, 264762, 264448, 264784, 264369, 264766, 18108357, 264768, 264687, 18108358, 264769, 55811957, 264690, 264691, 33657023, 264692, 18108362, 18108368, 264628, 264629, 18108374, 263978, 264634, 264635, 264636, 264637, 264638, 18108385, 264483, 264566, 264486, 264567 |
| 2221 | 88060927 (4441, 4442) | Novel Protein sim. GBank gi 3549154 (AC005625) - R27328_1 [Homo sapiens] | | | |
| 2222 | 84425892 (4443, 4444) | | | UNCLASSIFIED | 264908, 265020, 35695855 |
| 2223 | 95091849 (4445, 4446) | | | UNCLASSIFIED | 265010, 264685, 264690, 264693, 264628, 263974, 263976, 55811576, 264555, 264638, 83373044, 264483 |
| 2224 | 87388515 (4447, 4448) | Novel Protein sim. GBank gi 3876005 emb CAA84799 - (Z35719) cDNA EST EMBL:D87419 comes from this gene; cDNA EST EMBL:C13853 comes from this gene; cDNA EST EMBL:C11578 comes from this gene; cDNA EST yk234a7.3 comes from this gene; cDNA EST yk234a7.5 comes from this gene; cDNA ES.... | Contains protein domain (PF01958) - Domain of unknown function | UNCLASSIFIED | 264259, 264509, 56182435, 265006, 265008, 265009, 264757, 21908754, 18108351, 284693, 18108374, 18108385 |
| 2225 | 85749484 (4449, 4450) | Novel Protein sim. GBank gi 1255847 (U53338) - C05E11.1 gene product [Caenorhabditis elegans] | | transport | 22278994, 22278995, 22278999, 52644045, 264600, 265019, 21908765, 21908769 |
| 2226 | 86978953 (4451, 4452) | Novel Protein sim. GBank gi 4826524 emb CAB42852.1 - (AL049848) hypothetical protein [Homo sapiens] | | | 264259, 29331822, 29331824, 29331825, 29331827, 264508, 264908, 265007, 264691, 264634, 264486 |
| 2227 | 87721135 (4453, 4454) | | | UNCLASSIFIED | 22278999, 265006, 265008, 18108354, 29148628, 29148784, 27486261, 18108374, 264637, 18108384 |
| 2228 | 91227337 (4455, 4456) | Novel Protein sim. GBank gi 606976 (U16800) - ribonucleoprotein [Xenopus laevis] | Contains protein domain (PF00076) - RNA recognition motif. (a.k.a. RRM, RBD, or RNP domain) | dna_ma_bind | 264091, 264092, 264094, 29331822, 29331825, 66714117, 264693, 263972, 264639, 83373044, 264563 |
| 2229 | 88060931 (4457, 4458) | Novel Protein sim. GBank gi 3549155 (AC005625) - R27328_2 [Homo sapiens] | | UNCLASSIFIED | |

| | | | | |
|------|-----------------------|--|--|--|
| 2230 | 95342815 (4459, 4460) | Novel Protein sim. GBank gi 226154 pr 11412350A - DNA polymerase [Human adenovirus type 2] | UNCLASSIFIED | 264488, 264768, 52644507, 264769, 21906765, 21906766, 21906767, 21906769, 22278995, 35695917, 22278998, 22278997, 22278999, 265021, 264259, 52645129, 29331827, 264508, 264509, 264907, 18108370, 18108374, 35696423, 35695855, 265007, 264910, 264555, 33657402, 21906754, 18108387, 265010, 265018, 265019, 264760, 264288, 264567, 264563 |
| 2231 | 88060937 (4461, 4462) | Novel Protein sim. GBank gi 3549154 (AC005625) - R27328_1 [Homo sapiens] | UNCLASSIFIED | |
| 2232 | 87762581 (4463, 4464) | Novel Protein sim. GBank gi 5281316 gb AAD41476.1 AF13312 - (AF133124) transcription factor IIC63 [Homo sapiens] | transcription factor | 18108394, 56182575, 22278995, 35696286, 22278997, 22278999, 264259, 29331827, 35696052, 264907, 56182435, 265006, 265007, 265008, 264910, 264758, 55812038, 264603, 265018, 265019, 18108351, 264682, 264764, 264683, 264369, 264288, 264686, 264687, 264689, 21906765, 21906768, 21906767, 21906769, 29148629, 35695917, 264690, 52644150, 264691, 33657023, 264693, 18108370, 18108374, 55811576, 35695855, 264639, 18108385, 264564, 264906, 33657402, 265018, 264288, 264686, 265020, 264635, 18108385 |
| 2233 | 87752822 (4465, 4466) | Novel Protein sim. GBank gi 42487933 gb AAD13780 - (AF109377) IdIBp [Mus musculus] | | |
| 2234 | 87771817 (4467, 4468) | Novel Protein sim. GBank gi 1706559 sp P54352 EAS_DROME - ETHANOLAMINE KINASE (EASILY SHOCKED PROTEIN) | kinase | 56182435, 264369, 264688, 21906765, 265020, 264693, 264556, 56526486 |
| 2235 | 91012316 (4469, 4470) | Novel Protein sim. GBank gi 4972734 gb AAD34762.1 - (AF132174) unknown [Drosophila melanogaster] | UNCLASSIFIED | 22278997, 264563 |
| 2236 | 88003131 (4471, 4472) | Novel Protein sim. GBank gi 1082675 pr 1B53814 - p20 protein - human | WD domain, G-beta repeat Contains protein domain (PF00011) - eph Hsp20/alpha crystallin family | 264569, 264687, 264769, 265022, 264259, 60432049, 264691, 29331828, 60432289, 20281149, 264906, 264907, 264511, 265008, 265009, 264634, 264635, 264636, 264555, 264556, 264557, 264558, 60433356, 264595, 264559, 60433438, 60432113, 264761, 264762, 264763, 264764 |
| 2237 | 91012318 (4473, 4474) | Novel Protein sim. GBank gi 4972734 gb AAD34762.1 - (AF132174) unknown [Drosophila melanogaster] | kinase | 264488, 264769, 21906765, 21906766, 21906767, 21906768, 21906769, 22278995, 22278996, 22278997, 22278998, 265020, 265021, 22278999, 264259, 29331824, 29331826, 29331827, 35695763, 18108376, 35695855, 265007, 60432229, 33657402, 60433356, 60433438, 83373044, 18108385, 21906754, 18108387, 60432113, 22279000, 265019, 22279002, 264482, 18108351, 264288 |

| | | | | | |
|------|-----------------------|---|--|--------------|---|
| 2238 | 84998857 (4475, 4476) | | Contains protein domain (PF00286) - Viral coat protein | | 264509, 264907, 264629, 264634, 264564 |
| 2239 | 87798688 (4477, 4478) | | | | 29331825, 265009, 264369, 33657109, 18108370, 18108374, 264557, 264559 |
| 2240 | 94121471 (4479, 4480) | Novel Protein sim. GBank gij2882311 (AF051240) - probable ubiquitin-conjugating enzyme E2 [Picea mariana] | Contains protein domain (PF00179) - Ubiquitin-conjugating enzyme | ubiquitin | 264488, 65274572, 56182575, 35696286, 22278997, 22278999, 264259, 29331827, 35696052, 264508, 52644045, 56182435, 264511, 265007, 265008, 265009, 60433356, 60433438, 55812038, 21906754, 33657084, 55811386, 265018, 265019, 18108351, 264683, 264288, 264768, 264687, 264688, 264769, 21908765, 21908768, 21908769, 35695917, 265021, 265022, 60170615, 52644150, 33657023, 33657182, 33657349, 35695763, 18108370, 35696423, 35695855, 87168518, 22279000 |
| 2241 | 80091951 (4481, 4482) | | | UNCLASSIFIED | 264893, 264629 |
| 2242 | 91228075 (4483, 4484) | Novel Protein sim. GBank gij2494312 sp P70541 E2BG_RAT - TRANSLATION INITIATION FACTOR EIF-2B GAMMA SUBUNIT (EIF-2B GDP-GTP EXCHANGE FACTOR) | | synthase | 22278995, 22278996, 22278997, 22278998, 264259, 29331822, 29331824, 29331826, 29331827, 29331828, 264509, 265007, 265009, 264596, 21908754, 265010, 265011, 265017, 265018, 265019, 264448, 264369, 264288, 52644229, 21906765, 21906766, 21906767, 21906768, 21906769, 265020, 265021, 33657109, 27488282, 27486264, 18108374, 35695855, 264634, 264637, 56182323, 83373044, 56526486, 87168518, 264564 |
| 2243 | 78902026 (4485, 4486) | | | UNCLASSIFIED | 265008 |
| 2244 | 85723527 (4487, 4488) | Novel Protein sim. GBank gij2291143 (AF016417) - Similar to BZIP transcription factor [Caenorhabditis elegans] | | UNCLASSIFIED | 264604 |
| 2245 | 95318545 (4489, 4490) | Novel Protein sim. GBank gij470340 (U00043) - similar to beta-mannosyltransferase [Caenorhabditis elegans] | Contains protein domain (PF00534) - Glycosyl transferases group 1 | | 52645156, 22278995, 22278996, 22278997, 22278999, 29331822, 29331824, 29331827, 264907, 264512, 60433438, 264758, 21906754, 265011, 264603, 264764, 264687, 21906767, 21906768, 21906769, 55811957, 265022, 264691, 264629, 35696423, 264638, 18108387, 60432113, 22279000, 22279002, 264566 |

| | | | | | |
|------|-----------------------|---|--|--------------|--|
| 2246 | 94846710 (4481, 4492) | Novel Protein sim. GBank gl 496096 dbj BAA78326.1 - (ABD28068) activator of S phase Kinase (Homo sapiens) | Contains protein domain (PF00153) - Mitochondrial carrier proteins | transport | 65274572, 22278995, 35696286, 22278996, 22278997, 22278999, 264259, 35698052, 264106, 264905, 264907, 265006, 265007, 265008, 60433438, 33109954, 87168559, 265018, 265019, 264288, 21906765, 21906767, 21906768, 21906769, 55811957, 35695917, 265020, 265022, 27486284, 18108370, 18108374, 65274791, 35695855, 60432113 |
| 2247 | 87862542 (4493, 4494) | Novel Protein sim. GBank gl 854065 emb CAA58337 - (X83413) U88 [Human herpesvirus 6] | | UNCLASSIFIED | 52645156, 52646365, 52645080, 35696052, 33656970, 52648317, 33637084, 265017, 21906768, 21906769, 35695917, 33657109, 52645129, 33657182, 27486261, 27486282, 33657349, 27486265, 18108387 |
| 2248 | 95412996 (4495, 4496) | Novel Protein sim. GBank gl 4758502 ref NP_004123.1 pHABP - hyaluronan-binding protein 2 | Contains protein domain (PF00089) - Trypsin | calhepsin | 264488, 264259, 264907, 29331830, 264909, 265007, 265009, 264595, 21906754, 65274444, 264603, 265019, 264762, 264448, 264288, 264688, 21906766, 55811957, 265021, 264691, 18108374, 264634, 264635, 264636, 264555, 264638, 264557, 264558, 264559, 18108383, 83373044, 18108385, 264486 |
| 2249 | 94685662 (4497, 4498) | Novel Protein sim. GBank gl 4038461 (AF107772) - TcST11 [Trypanosoma cruzi] | Contains protein domain (PF00515) - TPR Domain | eph | 264766, 264628, 264636, 264637 |
| 2250 | 78827508 (4499, 4500) | Novel Protein sim. GBank gl 3738140 emb CAA21241 - (AL031852) valyl-trna synthetase, mitochondrial precursor [Schizosaccharomyces pombe] | | UNCLASSIFIED | 264908, 18108374 |
| 2251 | 87385863 (4501, 4502) | Novel Protein sim. GBank gl 3218467 emb CAA07090.1 - (AJ006529) putative phosphatase [Gallus gallus] | | UNCLASSIFIED | 264259, 35696052, 264508, 56182435, 265009, 264592, 264593, 264760, 264448, 264684, 264288, 264690, 264628, 55811576, 264555, 264556, 264557, 264558, 264559, 264568 |
| 2252 | 87735867 (4503, 4504) | Novel Protein sim. GBank gl 4929325 gb AAD33953.1 AF14531 - (AF145316) vacuolar proton pump delta polypeptide [Homo sapiens] | Contains protein domain (PF01813) - ATP synthase subunit D | synthase | 264092, 264094, 264259, 29331822, 66714117, 29331828, 264102, 264103, 264104, 264105, 264109, 264112, 264511, 265007, 60433356, 265010, 18108351, 21906767, 21906768, 264691, 263974, 263977, 264488, 264567 |
| 2253 | 91010703 (4505, 4506) | | | UNCLASSIFIED | 65274572, 265019 |

| | | | | | |
|------|-----------------------|--|---|--------------|--|
| 2254 | 95320031 (4507, 4508) | Novel Protein sim. GBank glj4502847[etNP_001271.1]pCIRB - cold inducible RNA- binding protein | Contains protein domain (PF000076) - RNA recognition motif (a.k.a. RRM, RBD, or RNP domain) | dna_rna_bind | 264569, 18108394, 18108398, 56182575, 56994075, 35696286, 22278999, 264094, 60432049, 264259, 29331822, 29331824, 29331825, 29331826, 60432289, 29331827, 29331828, 35696052, 264106, 264508, 264509, 264908, 264907, 29331830, 66712502, 264908, 264909, 264510, 265006, 264511, 265007, 265008, 265009, 60170831, 60432229, 60433356, 60433438, 264758, 85658542, 265010, 265011, 87168559, 265017, 265018, 265019, 264448, 264764, 264288, 264369, 264768, 264686, 264768, 264769, 21906765, 21906767, 55611957, 264691, 33657023, 264692, 18108362, 65274620, 263989, 264628, 18108370, 60431528, 263972, 264629, 18108372, 18108377, 18108379, 55811576, 35696423, 35695855, 264630, 264634, 264635, 264636, 264558, 263981, 264638, 56182323, 60170394, 264558, 18108381, 18108382, 83373044, 18108385, 87168518, 60432113, 22279002, 264482, 264564, 264565, 264486, 264587, 18108391 |
| 2255 | 81010546 (4509, 4510) | Novel Protein sim. GBank glj5541865[etmbjCAB51072.1] - (AL096858) hypothetical protein [Homo sapiens] | Contains protein domain (PF000076) - RNA recognition motif (a.k.a. RRM, RBD, or RNP domain) | strut | 65274572, 56182575, 22278997, 22278999, 264259, 29331822, 29331825, 29331826, 29331827, 29331828, 264508, 264905, 264906, 264907, 66712502, 264908, 56182435, 264510, 264511, 265008, 264593, 264595, 21906754, 33109954, 87168474, 265011, 265017, 265019, 264682, 264764, 264369, 264288, 264768, 264685, 264686, 264768, 21906765, 21906768, 21906769, 265020, 60170615, 52644150, 264690, 264692, 264693, 33657109, 33657349, 264632, 264636, 52644332, 56182323, 22279000, 22279002 |
| 2256 | 87020531 (4511, 4512) | Novel Protein sim. GBank glj3327174[etbjBAA31655] - (AB014580) KIAA0680 protein [Homo sapiens] | | UNCLASSIFIED | 264768, 264689, 18108374 |
| 2257 | 80088235 (4513, 4514) | | | UNCLASSIFIED | 22278996, 22278998, 264681, 21906765, 21906766, 264567 |
| 2258 | 88090516 (4515, 4516) | Novel Protein sim. GBank glj3025446 (AC004528) - R32184_2 [Homo sapiens] | Contains protein domain (PF000060) - Ligand-gated ion channel | misc_channel | 264808, 264592, 264764 |

| | | | | | |
|------|-----------------------|---|---|----------------|--|
| 2259 | 85384155 (4517, 4518) | Novel Protein sim. GBank gi 4884140 emb CAB43278.1 - (AL050110) hypothetical protein [Homo sapiens] | | UNCLASSIFIED | 18108398, 65274572, 56182575, 22278997, 22278998, 264259, 29331822, 29331827, 264905, 66712502, 264908, 264909, 56182435, 265007, 265008, 60432229, 33657084, 87168559, 18108351, 264448, 284683, 264288, 264369, 56181562, 265021, 60170615, 264690, 33657109, 60431528, 18108374, 52644332, 56182323, 18108385, 22279000, 22279002, 264482 |
| 2260 | 88084119 (4519, 4520) | Novel Protein sim. GBank gi 3080663 (AC004614) - similar to f-spondin proteins AB006086 (PID:g2529225) [Homo sapiens] | Contains protein domain (PF00090) - oxidase Thrombospondin type 1 domain | | 56182575, 265020, 264905, 264906, 264908, 35696423, 264511, 264635, 55812038, 264758, 265018, 265019, 264605, 264760, 264563 |
| 2261 | 88074157 (4521, 4522) | Novel Protein sim. GBank gi 3334526 emb CAA16138 - (AL021308) predicted using FGENEH [Homo sapiens] | | UNCLASSIFIED | |
| 2262 | 91639282 (4523, 4524) | Novel Protein sim. GBank gi 4877759 gb AAD31421.1 AF12444 - (AF124440) MAGE tumor antigen D1 [Homo sapiens] | | | 56182575, 22278999, 29331822, 29331825, 60432289, 29331827, 35696052, 264508, 66712502, 52644045, 56182435, 265008, 265008, 265009, 60433356, 55812038, 265010, 265017, 265019, 264288, 264369, 21906765, 21906767, 55811957, 35695917, 52644150, 33657023, 33657109, 55811576, 65274791, 56182323 |
| 2263 | 87602495 (4525, 4526) | Novel Protein sim. GBank gi 3341697 (AC003672) - hypothetical protein [Arabidopsis thaliana] | | | 22278994, 22278997, 264907, 264828, 52644150, 18108381, 264893, 18108374 |
| 2264 | 87756525 (4527, 4528) | Novel Protein sim. GBank gi 657601 (U66220) - unknown [Nannocystis exedens] | | UNCLASSIFIED | 264686, 264488, 264768, 264769, 264691, 264508, 264905, 264509, 264906, 264907, 264908, 264909, 35695855, 264510, 264511, 264512, 265007, 265009, 264638, 264639, 264757, 264758, 18108385, 265011, 264760, 264564, 264565, 264764, 264566, 264766, 264766 |
| 2265 | 86918663 (4529, 4530) | Novel Protein sim. GBank gi 477072 pir A48018 - mucin 7 precursor, salivary - human | Contains protein domain (PF00096) - Zinc finger, C2H2 type | UNCLASSIFIED | 264689, 264910, 264764 |
| 2266 | 87773458 (4531, 4532) | Novel Protein sim. GBank gi 3150479 (AF067212) - partial CDS [Caenorhabditis elegans] | Contains protein domain (PF01305) - Ribosomal protein L15 amino terminal region | ribosomal prot | 22278995, 22278997, 22278999, 264259, 265008, 265007, 265009, 60433438, 21906754, 265010, 265011, 265017, 264448, 264683, 264288, 264689, 21906765, 21906768, 35695917, 265021, 18108374, 264638, 22279000, 22279002, 264566, 264487 |

| | | | | | |
|------|-----------------------|---|--|--------------|--|
| 2267 | 87395838 (4533, 4534) | Novel Protein sim. GBank gij3560229[emb]CAA20697.1] - (AL031530) hypothetical protein [Schizosaccharomyces pombe] | | UNCLASSIFIED | 35696286, 264259, 29331824, 29331825, 35696052, 29331828, 264905, 264509, 264907, 264908, 264909, 264512, 265009, 264910, 264593, 33657402, 265010, 265018, 264762, 264448, 264288, 264369, 264768, 52644229, 35695917, 264691, 33657023, 18108362, 33657109, 35696423, 264634, 18108381, 87168518, 264566 |
| 2268 | 85693687 (4535, 4536) | Novel Protein sim. GBank gij728632[sp]P39189[ALU2_SB WARNING ENTRY IIII | cadherin | | 264488, 264259, 264509, 264595, 265010, 265017, 264766, 18108385, 264486 |
| 2269 | 88177877 (4537, 4538) | Novel Protein sim. GBank gij103418[pir][S17885 - Tcd37 protein - fruit fly (Drosophila melanogaster) | UNCLASSIFIED | | 56182575, 60432049, 265007, 265009, 264591, 87169559, 264605, 18108351, 21906764, 265020, 264629, 60431528, 264638, 18108385, 18108387, 60432113 |
| 2270 | 80410327 (4539, 4540) | | | | 264763 |
| 2271 | 91010392 (4541, 4542) | | cyto450 | | 264909, 56182435, 265008, 55812038, 55811957, 33657023, 264693, 33657109, 55810784, 55811576, 56182323 |
| 2272 | 84208220 (4543, 4544) | | UNCLASSIFIED | | 264905, 264908 |
| 2273 | 95014271 (4545, 4546) | Novel Protein sim. GBank gij4176370 (AC005058) - similar to calcium-independent phospholipase A2; similar to AC004392 (PID:g3367519) [Homo sapiens] | Contains protein domain (PF00462) - Glutaredoxin | | 52645156, 22278996, 22278999, 60432049, 264259, 29331822, 29331824, 29331825, 29331826, 29331827, 35696052, 264909, 265006, 264593, 60433438, 21906754, 265018, 264689, 21906765, 21906766, 21906767, 21906769, 265021, 265022, 60170615, 264691, 33657023, 264693, 33657109, 27486264, 18108376, 35696423, 35695855, 264630, 52644332, 264558, 56182323, 22279002 |
| 2274 | 91640217 (4547, 4548) | Novel Protein sim. GBank gij1480112[emb]CAA67961] - (X99842) HP1-BP74 protein [Mus musculus] | histone linker histone H1 and H5 family | | 52645156, 22278997, 22278999, 52645080, 29331824, 29331825, 29331826, 29331827, 29331828, 264905, 264908, 52644045, 264511, 265008, 265009, 60170831, 264591, 21906754, 33109954, 265011, 265018, 18108351, 264448, 264288, 264684, 264766, 21906765, 21906766, 21906767, 21906768, 52644150, 264693, 18108384, 35695763, 18108374, 35698423, 264634, 264557, 264638, 52644332, 83373044, 18108385, 56526486, 87168518, 22279002 |
| 2275 | 88082501 (4549, 4550) | Novel Protein sim. GBank gij3165406 (AC004755) - [os37502_2 [Homo sapiens] | Contains protein domain (PF00122) - E1-E2 ATPase | transport | |
| 2276 | 11287447 (4551, 4552) | | UNCLASSIFIED | | 264555, 264556 |

| | | | | | |
|------|-----------------------|---|--|--------------|--|
| 2277 | 88084123 (4553, 4554) | Novel Protein sim. GBank gii2880079 (AC004142) - similar to murine leucine-rich repeat protein; possible role in neural development by protein-protein interactions; 93% similarity to D48802 (PID:g1369906) [Homo sapiens] | Contains protein domain (PF00560) - Leucine Rich Repeat | glycoprotein | 22278999, 35696052, 265008, 265019, 264369, 265020, 265022, 55810764, 264404, 22278002 |
| 2278 | 84133079 (4555, 4556) | Novel Protein sim. GBank gii2618702 (AC002510) - unknown protein [Arabidopsis thaliana] | | UNCLASSIFIED | 18108394, 22278997, 29331826, 60433356, 60433438, 21906754, 285018, 33657023, 264639, 83373044, 264565 |
| 2279 | 80419375 (4557, 4558) | Novel Protein sim. GBank gii119714 [sp]P13983 [EXTN_TOBAC - EXTENSIN PRECURSOR (CELL WALL HYDROXYPROLINE-RICH GLYCOPROTEIN)] | | UNCLASSIFIED | 264766, 264565 |
| 2280 | 84239723 (4559, 4560) | | | | 265008, 33109954, 265010, 265019, 265020 |
| 2281 | 95293048 (4561, 4562) | Novel Protein sim. GBank gii4240299 [dbj]BAA74928.1] - (AB020712) KIAA0905 protein [Homo sapiens] | Contains protein domain (PF00400) - WD domain, G-beta repeat | transport | 264092, 264259, 29331822, 29331824, 29331826, 35696052, 264107, 264908, 264909, 52844045, 265006, 33657402, 60433356, 264758, 265011, 285019, 264881, 264683, 264684, 264686, 21908765, 21908767, 21906768, 21906769, 60170615, 264690, 52644150, 18108362, 264692, 18108368, 18108374, 263978, 264631, 18108381, 264559, 18108385, 56528486, 22279000, 264566, 264567 |
| 2282 | 87602829 (4563, 4564) | Novel Protein sim. GBank gii1537070 (U63840) - nucleoporin p54 [Rattus norvegicus] | | UNCLASSIFIED | 264488, 264259, 29331822, 29331824, 29331827, 29331828, 29331830, 33657402, 60433438, 87168474, 265019, 18108351, 21906767, 21906769, 55811957, 33657023, 52645129, 33657109, 33657182, 27486282, 263972, 55811576, 87168518, 20281169 |
| 2283 | 95362388 (4565, 4566) | Novel Protein sim. GBank gii249572 [sp]Q92556 [Y281_HUMAN - HYPOTHETICAL PROTEIN KIAA0281 (HA6725)] | | UNCLASSIFIED | 60424179, 56182575, 22278994, 35696286, 22278997, 22278999, 29331822, 29331824, 56182181, 29331825, 29331827, 35696052, 29146499, 264905, 66712502, 264908, 265007, 265009, 60432229, 264593, 60431735, 60433356, 33109954, 33657084, 55811388, 87168474, 265010, 265011, 265018, 265019, 55811150, 264683, 264369, 264288, 264688, 21906765, 21908767, 21908768, 29148627, 21908769, 55811957, 265020, 265022, 33657182, 27486261, 18108370, 264628, 18108374, 55810764, 18108379, 55811576, 35696423, 35695855, 264630, 60431850, 263981, 18108382, 83373044, 18108385, 18108387, 60432113, 22279000, 264482, 264567 |

| | | | | | |
|------|-----------------------|--|--------------|--|---|
| 2284 | 95414955 (4567, 4568) | Novel Protein sim. GBank gi 2498797 sp Q64311 PNAD_MOUSE - PROTEIN N- TERMINAL ASPARAGINE AMIDOHYDROLASE (PROTEIN NH2-TERMINAL ASPARAGINE DEAMIDASE) (NTN- AMIDASE) (PNAD) (PROTEIN NH2-TERMINAL ASPARAGINE AMIDOHYDROLASE) (PNAAD) | | | 60424179, 52644507, 18108394, 52646842, 22278994, 35696286, 22278996, 22278997, 22278999, 264259, 60432049, 29331822, 29331824, 29331825, 60432289, 29331826, 29331827, 35696052, 29331830, 52644045, 56182435, 33657402, 60433438, 33109954, 21906754, 85658542, 87168559, 265018, 265019, 55811150, 264682, 264369, 21906765, 21906766, 21908767, 21908768, 21906769, 55811957, 35695917, 265020, 265021, 60170815, 33657023, 33657182, 27486262, 27486264, 27486265, 18108376, 55810764, 35698423, 35695855, 60431850, 87168518, 60432113, 264482, 264584 |
| 2285 | 87781484 (4569, 4570) | Novel Protein sim. GBank gi 3342234 (U83909) - nuclear antigen EBNA-1 [Cercopithecine herpesvirus 15] | collagen | | 35696052, 264905, 264907, 264908, 264909, 264512, 265009, 264910, 264595, 264760, 18108351, 264682, 264763, 264685, 264766, 264686, 264768, 264693, 264629, 35695855, 264631, 264634 |
| 2286 | 87737825 (4571, 4572) | Novel Protein sim. GBank gi 3873414 (U00043) - similar to D. melanogaster trithorax protein [Caenorhabditis elegans] | kinase | | 35696286, 58182435, 60170831, 264591, 60432229, 264592, 264593, 264594, 264595, 55812038, 264596, 87168474, 35695917, 264692, 55811578, 264555, 264557 |
| 2287 | 82986696 (4573, 4574) | Novel Protein sim. GBank gi 630905 pir S42731 - collagen alpha 1 chain - sea urchin (Hemicentrotus pulcherrimus) (fragment) | UNCLASSIFIED | | 264682 |
| 2288 | 94133083 (4575, 4576) | Novel Protein sim. GBank gi 728832 sp P39189 ALU2_HUMAN - !!!!! ALU SUBFAMILY SB WARNING ENTRY !!!!! | tm7 | | 22278995, 56994075, 22278997, 22278999, 264259, 60432289, 264508, 264512, 265008, 33657402, 265017, 265018, 265019, 18108351, 264448, 21906765, 21906766, 21906767, 21906768, 33657023, 264557, 22279000, 22279002 |
| 2289 | 88084133 (4577, 4578) | Novel Protein sim. GBank gi 2887497 (AC004144) - R34001_1 [Homo sapiens] | kinase | | |
| 2280 | 88084137 (4579, 4580) | Novel Protein sim. GBank gi 2887497 (AC004144) - R34001_1 [Homo sapiens] | UNCLASSIFIED | | 264683 |
| 2291 | 84285281 (4581, 4582) | Novel Protein sim. GBank gi 3253120 (AC005175) - R31449_3 [Homo sapiens] | strud | | 18108394, 264907, 265006, 265009, 33109954, 52646317, 265010, 18108351, 264681, 264686, 264692, 18108370, 18108374, 18108385 |

| | | | | | |
|------|-----------------------|--|--|------------------|---|
| 2282 | 94328834 (4583, 4584) | Novel Protein sim. GBank gj14803672[emb]CAB42643.1]- (AJ133769) nuclear transport receptor [Homo sapiens] | | UNCLASSIFIED | 56182575, 35696286, 56994075, 29331824, 29331825, 35696052, 56182435, 60433438, 55812038, 33109954, 87168474, 87168559, 265018, 18108351, 264763, 264448, 264369, 264288, 58181562, 264769, 21906765, 21906766, 21906767, 21906769, 265021, 265022, 33657023, 264693, 65274620, 33657109, 27486264, 264629, 55810764, 55811576, 35695855, 58182323, 56526486, 87168518, 22279000, 264567 |
| 2283 | 87759213 (4585, 4586) | Novel Protein sim. GBank gj3252981 (AF068921) - Ras- binding protein SUR-8 [Mus musculus] | Contains protein domain (PF00560) - Leucine Rich Repeat | siRud | 264488, 18108397, 35696286, 264092, 264259, 29331822, 29331826, 264906, 264908, 264511, 264512, 265009, 264910, 18108351, 264764, 264369, 264288, 264685, 264766, 265020, 265022, 264534, 35696423, 264631, 264637, 18108381, 56182323, 264639, 18108385, 264404, 264563, 264565 22278997, 22278998, 22278999, 264259, 29331822, 60432289, 29331828, 35696052, 265018, 264684, 264288, 264686, 21906765, 21906766, 21906767, 21906768, 21906769, 265020, 265021, 264692, 33657109, 18108376, 35696423, 35695855, 264634, 22279000, 22279002, 264563, 264486 22278996, 60432289, 264682, 264683, 264689, 18108374 |
| 2284 | 86693580 (4587, 4588) | Novel Protein sim. GBank gj2062680 (U88964) - HEM45 [Homo sapiens] | Contains protein domain (PF00929) - Exonuclease | nuclease | 263974, 263978 |
| 2285 | 95312200 (4589, 4590) | | | UNCLASSIFIED | 264488, 65274572, 56182575, 22278997, 22278999, 264259, 29331822, 29331824, 29331828, 29331828, 35696052, 264907, 264908, 52644045, 56182435, 264112, 265006, 265007, 264910, 265009, 60433356, 33657402, 264595, 55812038, 21906754, 265011, 265018, 265019, 264448, 264764, 264288, 264766, 21906765, 21906767, 21906768, 21906769, 55811957, 265020, 265021, 264690, 264691, 33657023, 264692, 264693, 33657109, 55811576, 56182323, 60170394, 83373044, 18108385, 56526486, 264564, 264486 |
| 2286 | 80030781 (4591, 4592) | | | | |
| 2287 | 94321251 (4593, 4594) | Novel Protein sim. GBank gj5689501[dbj]BAA83034.1]- (AB029005) KIAA1082 protein [Homo sapiens] | | transcriptfactor | |

| | | | | | |
|------|-----------------------|--|---|--------------|---|
| 2298 | 95312207 (4595, 4598) | Novel Protein sim. GBank gi 3875051 emb CAB02849 - (Z81050) predicted using Genefinder; similar to collagen; cDNA EST EMBL:D85584 comes from this gene; cDNA EST EMBL:D89046 comes from this gene; cDNA EST yk366b12.3 comes from this gene; cDNA EST yk366b12.5 comes from this gene ... | | collagen | 6042478, 56181686, 22278995, 35686286, 22278996, 22278998, 22278999, 264490, 264259, 29331822, 29331824, 66714117, 60424269, 35686052, 29331828, 66712502, 56182435, 264510, 265006, 60433438, 21906754, 33109854, 55811386, 285010, 265018, 55811150, 264762, 18108351, 264682, 264683, 264288, 264684, 264686, 264688, 56181562, 264689, 21906766, 21906767, 29148629, 55811957, 29148784, 35695917, 265020, 18108362, 33657023, 18108364, 33657109, 60431602, 18108370, 60431528, 18108374, 55810764, 35698423, 35695855, 264630, 264634, 60431850, 18108380, 56182323, 264558, 83373044, 18108385, 60432113, 22278000, 264482, 264567, 264486 |
| 2299 | 80193720 (4597, 4598) | | | UNCLASSIFIED | 264369 |
| 2300 | 84124346 (4599, 4600) | Novel Protein sim. GBank gi 2443886 (AC002284) - Unknown protein [Arabidopsis thaliana] | | | 264488, 22278996, 22278999, 264259, 29331824, 66714117, 35686052, 264509, 264905, 264906, 264907, 264908, 264909, 265008, 264910, 265009, 264758, 265010, 87168559, 264600, 265018, 264760, 264762, 18108351, 264764, 264766, 264768, 264769, 21906766, 21906767, 35695917, 265021, 264691, 33657023, 35695763, 18108370, 18108374, 35698423, 35695855, 264631, 264636, 264638, 18108385, 22279002, 264563 |
| 2301 | 91235725 (4601, 4602) | Novel Protein sim. GBank gi 2143637 pir 84505 - calcium-dependent actin-binding protein - rat | | struct | 264908, 264758, 265017, 21906765, 83373044, 264563 |
| 2302 | 88084141 (4603, 4604) | Novel Protein sim. GBank gi 2887487 (AC004144) - R34001.1 [Homo sapiens] | | UNCLASSIFIED | 57644045, 265019, 264288, 33657023, 18108370, 18108385 |
| 2303 | 94141439 (4605, 4606) | Novel Protein sim. GBank gi 4884194 emb CAB43220.1 - (AL049946) hypothetical protein [Homo sapiens] | Contains protein domain (PF00047) - struct Immunoglobulin domain | | 264259, 60432049, 264907, 264909, 264910, 60432229, 33657402, 265011, 265018, 264762, 264448, 264769, 264637, 264638, 83373044, 264486 |
| 2304 | 94840434 (4607, 4608) | Novel Protein sim. GBank gi 2494162 sp Q10005 YRY1_CAEEL - HYPOTHETICAL 39.9 KD PROTEIN T15H9.1 IN CHROMOSOME II PRECURSOR | | UNCLASSIFIED | 264259, 29331824, 21906767, 33657182, 33657349 |
| 2305 | 90935911 (4609, 4610) | Novel Protein sim. GBank gi 4872686 gb AAD34738.1 - (AF132150) unknown [Drosophila melanogaster] | | | 83274572, 22278996, 264908, 265006, 21906769, 264691, 264488 |

| | | | | | |
|------|-----------------------|---|---|------------------|--|
| 2308 | 95334940 (4611, 4612) | Novel Protein sim. GBank gi 4929565 gb AAD34043.1 AF15180 - (AF151806) CGI-48 protein [Homo sapiens] | Contains protein domain (PF00400) - WD domain, G-beta repeat | kinasereceptor | 264488, 22278995, 22278996, 35696286, 22278997, 22278998, 22278999, 264259, 29331822, 29331824, 29331825, 29331826, 29331827, 29331828, 35696052, 264508, 264905, 264907, 29331830, 264908, 264909, 264510, 265006, 264511, 265007, 264512, 265008, 265009, 33657402, 21906754, 85658542, 265010, 265011, 264601, 265017, 265018, 264604, 265019, 18108351, 264448, 264288, 264766, 264769, 21906765, 21906766, 21906767, 21906768, 29148629, 28148784, 35695917, 265020, 265021, 265022, 33657023, 264692, 18108370, 18108374, 18108376, 35696423, 35695855, 264630, 264634, 264635, 264636, 264637, 264638, 264639, 18108382, 18108385, 18108387, 264563, 264566, 264486 |
| 2307 | 79415283 (4613, 4614) | Novel Protein sim. GBank gi 4758732 ref NP_004522.1 pMOCS - molybdenum cofactor synthesis 2 | | UNCLASSIFIED | 264828 |
| 2308 | 87608409 (4615, 4616) | Novel Protein sim. GBank gi 4758732 ref NP_004522.1 pMOCS - molybdenum cofactor synthesis 2 | | synthase | 35696286, 264259, 29331822, 29331824, 264112, 264512, 264757, 21906754, 264288, 264690, 27486284, 264631, 264634, 264404 |
| 2309 | 95357218 (4617, 4618) | Novel Protein sim. GBank gi 3878058 emb CAB17070 - (Z99942) cDNA EST EMBL:D73444 comes from this gene; cDNA EST EMBL:D70905 comes from this gene; cDNA EST EMBL:D72208 comes from this gene; cDNA EST EMBL:D75030 comes from this gene; cDNA EST EMBL:D72944 comes from this gene; cDN... | | glycoprotein | 18108397, 22278996, 22278997, 22278998, 22278999, 60432049, 29331822, 29331826, 60432289, 66712502, 60432228, 60433356, 60433438, 65274444, 265010, 264600, 264681, 264448, 264683, 264288, 21906766, 21906768, 265020, 264691, 264692, 264693, 65274620, 65274791 |
| 2310 | 79601668 (4619, 4620) | Novel Protein sim. GBank gi 2137337 pir J148281 - gene mCBP protein - mouse | Contains protein domain (PF00013) - KH domain | UNCLASSIFIED | 264508 |
| 2311 | 87721189 (4621, 4622) | Novel Protein sim. GBank gi 2137337 pir J148281 - gene mCBP protein - mouse | | transcriptfactor | 18108397, 56182575, 22278996, 56994075, 264259, 29331824, 29331827, 264508, 264907, 56182435, 264510, 264511, 265006, 264512, 265007, 265008, 265009, 60433438, 33109954, 265010, 265011, 264603, 265017, 18108351, 264762, 264683, 264288, 264369, 264686, 33657023, 20281149, 20281069, 264628, 263972, 55811576, 35696423, 20281071, 264632, 264636, 18108385, 18108387, 87168518, 22278990, 264563, 264486 |

| | | | | | |
|------|-----------------------|---|---|--|--|
| 2312 | 87549681 (4623, 4624) | Novel Protein sim. GBank gij2911264 (AC002550) - Unknown gene product [Homo sapiens] | | | 56182575, 56994075, 35696286, 22278996, 22278997, 22278999, 264259, 29331822, 29331824, 29331825, 29331826, 29331827, 52644045, 60432229, 60433356, 55812038, 33109954, 21906754, 87168474, 265018, 18108351, 264288, 52644229, 21906765, 21906766, 21906767, 21906768, 21906769, 265022, 60170815, 33657023, 27486261, 27486264, 35696423, 35695855, 18108385, 22279000, 22279002 263981 |
| 2313 | 80042533 (4625, 4626) | Novel Protein sim. GBank gij3043626(dbj BAA25477) - (AB011123) KIAA0551 protein [Homo sapiens] | | | |
| 2314 | 94313401 (4627, 4628) | Novel Protein sim. GBank gij5596714(emb CAB51401.1) - (AL035398) dJ796117.2 (CGI-51) [Homo sapiens] | UNCLASSIFIED | | 52644507, 52646365, 52646842, 22278996, 22278997, 22278999, 264259, 52645080, 29331822, 29331824, 60432289, 29331826, 29331827, 29331828, 35696052, 33656970, 52644045, 265008, 264593, 60433356, 60433438, 264758, 33109954, 265010, 265017, 265018, 265019, 264288, 264369, 21906765, 21906766, 21906768, 35695917, 52644150, 33657023, 33657109, 52645129, 33657349, 35695763, 18108374, 35696423, 35695855, 52644332, 22279000, 22279002, 264553, 264567 |
| 2315 | 80430119 (4629, 4630) | | UNCLASSIFIED | | 264905, 264906, 264767, 264768, 264693, 55811576, 264635, 56182323, 18108385 |
| 2316 | 94312191 (4631, 4632) | Novel Protein sim. GBank gij5531827(jb AAD44488.1) - (AF078856) p47 [Homo sapiens] | Contains protein domain (PF00789) - glycoprotein UBX domain | | 52644507, 52645158, 52646365, 22278994, 22278995, 22278996, 56994075, 35696288, 22278997, 22278998, 22278999, 264259, 29331822, 52645080, 29331824, 29331825, 29331826, 29331827, 35696052, 29331828, 264906, 264907, 29331830, 52644045, 56182435, 264511, 265007, 265008, 265009, 60170831, 60433438, 21906754, 52646317, 33109954, 33657084, 52644296, 87168474, 265010, 87168559, 265017, 265018, 265019, 264681, 264763, 264448, 264683, 264369, 52644229, 21906764, 264689, 21906765, 21906766, 21906767, 21906768, 21906769, 35695917, 265020, 265021, 52644150, 33657023, 18108362, 52645129, 33657182, 33657349, 35695763, 18108370, 18108376, 35696423, 35695855, 264631, 264556, 52644332, 83373044, 18108385, 18108387, 87168518, 60432113, 22279000, 264566, 264567 |

| | | | | | |
|------|-----------------------|--|---|---------------|--|
| 2317 | 87020571 (4633, 4634) | | | UNCLASSIFIED | 22278998, 60432049, 264910, 60432229, 264686, 264687, 264688, 264689, 264558, 18108385 |
| 2318 | 79959979 (4635, 4636) | | | UNCLASSIFIED | 264488, 264569, 18108396, 52646365, 22278994, 22278995, 22278996, 58994075, 35696286, 22278997, 22278998, 264259, 52645080, 29331825, 29331826, 29331827, 29331828, 29331830, 56182435, 60170831, 60432229, 60431735, 33657402, 21908754, 52644296, 87168474, 265011, 87168559, 265017, 265018, 265019, 18108351, 264448, 18108354, 264288, 264369, 52644229, 21908764, 21908765, 21908766, 21908767, 21908768, 21908769, 265021, 265022, 52644150, 33657023, 52645129, 33657109, 27486284, 33657349, 35695763, 18108370, 18108376, 18108379, 35696423, 264558, 83373044, 18108385, 56526486, 87168518, 264564, 264565, 264566 |
| 2320 | 91622426 (4638, 4640) | Novel Protein sim. GBank gi 728837 sp P39194 ALU7_HUMAN - IIII ALU SUBFAMILY SQ WARNING ENTRY IIII | | kinase | 22278994, 60432049, 60432289, 29331827, 264511, 265008, 52646317, 265017, 265019, 21908765, 18108372, 18108387, 22279002 |
| 2321 | 94320377 (4641, 4642) | Novel Protein sim. GBank gi 3873837 emb CAB02700 (Z81029) Similarity to S.pombe hypothetical protein C1D4.09C (SW:Q10154); cDNA EST EMBL: T00543 comes from this gene; cDNA EST EMBL: T01062 comes from this gene; cDNA EST EMBL: T01321 comes from this gene; cDNA EST EMBL: T02288 com... | | UNCLASSIFIED | 264488, 264687, 18108394, 264689, 21908765, 18108397, 18108398, 21908767, 21908768, 65274791, 22278995, 35695855, 22278998, 265021, 265022, 264510, 265006, 264511, 264512, 265008, 60170815, 264555, 264636, 264556, 18108361, 264259, 60432229, 33657023, 264557, 264558, 264693, 60433356, 264559, 60433438, 29331824, 18108365, 18108348, 18108384, 29331825, 18108385, 33109954, 29331827, 56526486, 29146499, 265011, 60432113, 265017, 265018, 264508, 264563, 264482, 264509, 18108351, 264448, 264907, 264682, 18108370, 264683, 264908, 264288, 264909, 18108354, 264486, 264567 |
| 2322 | 87803165 (4643, 4644) | Novel Protein sim. GBank gi 5678957 emb CAB51685.1 (AL109630) BACR7A4.y [Drosophila melanogaster] | Contains protein domain (PF00106) - short chain dehydrogenase | dehydrogenase | 22278996, 264907, 264511, 264757, 18108351, 264768, 264638 |

| | | | | |
|------|-----------------------|--|--|--|
| 2323 | 94840445 (4845, 4846) | Novel Protein sim. GBank gi 2494162 sp Q10005 YRY1_CAEEL - HYPOTHETICAL 39.9 KD PROTEIN T15H9.1 IN CHROMOSOME II PRECURSOR | Contains protein domain (PF00226) - eph DnaJ domain | 22278994, 22278995, 22278997, 60432049, 264259, 29331822, 33656970, 264509, 56182435, 264511, 265008, 60433356, 60433438, 55812038, 33109954, 21906754, 85656542, 87168474, 285011, 87168559, 265017, 265019, 264760, 264681, 18108351, 264369, 264288, 18108355, 264687, 264688, 21906765, 21906767, 21906768, 55811957, 35695917, 265021, 33657023, 18108362, 27486262, 55811576, 264631, 264555, 83373044, 87168518, 60432113, 22279002 264592, 264593, 265020 |
| 2324 | 86633607 (4647, 4648) | Novel Protein sim. GBank gi 5419865 emb CAB46377.1 - (AL096732) hypothetical protein [Homo sapiens] | ATPase associated | 265020 |
| 2325 | 88165074 (4649, 4650) | Novel Protein sim. GBank gi 231985 sp P29981 CP4C_BLADI - CYTOCHROME P450 4C1 (CYP1VC1) | Contains protein domain (PF00067) - cyto450 Cytochrome P450 | 265006, 264759, 35695855, 56182323 |
| 2326 | 84390962 (4651, 4652) | Novel Protein sim. GBank gi 4240227 dbj BAA74892.1 - (AB020676) KIAA0869 protein [Homo sapiens] | UNCLASSIFIED | |
| 2327 | 88081648 (4653, 4654) | Novel Protein sim. GBank gi 1245105 U46463.1 - glutamine repeat protein-1 [Mus musculus] | UNCLASSIFIED | 264259, 264508, 264905, 264906, 264907, 264908, 265007, 264512, 264910, 264756, 265010, 264766, 264768, 264769, 33657023, 264693, 264628, 264631, 264634, 264638, 264639, 264486 |
| 2328 | 83388428 (4655, 4656) | Novel Protein sim. GBank gi 169343 sp P42209 DIF6_MOUSE - DIFF6 PROTEIN | UNCLASSIFIED | 60433438, 264595, 265017, 264766, 264692, 264629, 264635, 264636, 264638, 56182323, 60432113, 264566 |
| 2329 | 87604478 (4657, 4658) | Novel Protein sim. GBank gi 169343 sp P42209 DIF6_MOUSE - DIFF6 PROTEIN | UNCLASSIFIED | 265017, 264685, 60432113, 264088 |
| 2330 | 87335396 (4659, 4660) | Novel Protein sim. GBank gi 5679136 gb AA046874.1 AF16093 - (AF160934) BcDNA LD14189 [Drosophila melanogaster] | transport | 265009 |
| 2331 | 86990483 (4661, 4662) | Novel Protein sim. GBank gi 5679136 gb AA046874.1 AF16093 - (AF160934) BcDNA LD14189 [Drosophila melanogaster] | ATPase associated | 35696286, 22278998, 29331824, 60424269, 265006, 265008, 265018, 264448, 264764, 21906765, 35695917, 35695855, 264636, 22279000, 264566 |
| 2332 | 87784182 (4663, 4664) | Novel Protein sim. GBank gi 2104452 emb CAB08779 - (Z95397) unknown [Schizosaccharomyces pombe] | UNCLASSIFIED | 56182575, 56994075, 29331826, 29331828, 264107, 33657402, 87168559, 264683, 35695917, 265021, 33657023, 263976 |
| 2333 | 88206958 (4665, 4666) | Novel Protein sim. GBank gi 3879985 emb CAA92691.1 - (Z68318) cDNA EST CEMSD62F comes from this gene; cDNA EST EMBL:C07930 comes from this gene; cDNA EST EMBL:C08493 comes from this gene; cDNA EST yk415e8.3 comes from this gene; cDNA EST yk415e8.5 comes from this gene; cDNA EST... | UNCLASSIFIED | |
| 2334 | 84319788 (4667, 4668) | Novel Protein sim. GBank gi 4966270 gb AAB52261.2 - (U97002) similar to acyl-CoA dehydrogenases and epoxide hydrolases; Pfam domain PF00441 (Acyl-CoA_dh), Score=57.4, E-value=1.7e-16, N=2; contains similarity to Pfam domain PF00702 (Hydrolase), Score=57.4, E- value=1e-13, N=1 [C... | Contains protein domain (PF00441) - Acyl-CoA dehydrogenase | 56182575, 29331825, 21906768, 264636, 83373044 |

| | | | | | |
|------|-----------------------|--|---|--------------|--|
| 2335 | 80046103 (4669, 4670) | Novel Protein sim. GBank gi 32833350 (AF062378) - calmodulin-binding protein SHA1 [Mus musculus] | Contains protein domain (PF00612) - IQ calmodulin-binding motif | struct | 18108351, 21906769, 264555 |
| 2336 | 95196121 (4671, 4672) | Novel Protein sim. GBank gi 1929056jemb CAA72805j - (Y12090) putative 3,4-dihydroxy-2-butanone kinase [Lycopersicon esculentum] | | kinase | 264907, 35695917, 18108379 |
| 2337 | 95345810 (4673, 4674) | Novel Protein sim. GBank gi 4495063jemb CA839181.1j - (Z85986) dJ108K1.1.3 (similar to yeast suppressor protein SRP40) [Homo sapiens] | | UNCLASSIFIED | 35696286, 22278999, 56182181, 29331825, 60424269, 56182435, 33657402, 55812038, 55811386, 265017, 265018, 265019, 21906768, 35695917, 284691, 33657023, 33657109, 263972, 35696423, 35695855, 60432113 |
| 2338 | 87634045 (4675, 4676) | Novel Protein sim. GBank gi 2224689jdbj BAA20829j - (AB002372) KIAA0374 [Homo sapiens] | Contains protein domain (PF00323) - Mammalian defensin | UNCLASSIFIED | 18108394, 29331822, 66714117, 60432289, 264906, 56182435, 265009, 60433438, 264596, 265010, 265019, 18108354, 264288, 264368, 55811957, 265021, 33657023, 263976, 55811576, 264632, 56182323, 264639 |
| 2339 | 85663319 (4677, 4678) | Novel Protein sim. GBank gi 3873550jemb CAA22127j - (AL033534) serine-rich protein [Schizosaccharomyces pombe] | | UNCLASSIFIED | 35696286, 264592, 264369, 264691, 264558 |
| 2340 | 90937716 (4679, 4680) | | | | 65274572, 22278994, 35696286, 22278997, 22278999, 264259, 29331822, 60432289, 29331826, 29331830, 265009, 33657402, 33109954, 265017, 265018, 264766, 264685, 21906769, 35695917, 284691, 264692, 35696423, 87168518, 22279000 |
| 2341 | 8775281 (4681, 4682) | Novel Protein sim. GBank gi 3874563jemb CA802797j - (Z81042) similar to Yeast hypothetical protein Y6Y6 like; cDNA EST yk206h5.3 comes from this gene; cDNA EST yk206h5.5 comes from this gene; cDNA EST yk303h1.3 comes from this gene; cDNA EST yk303h1.5 comes from this gene; cDNA ... | Contains protein domain (PF00400) - WD domain, G-beta repeat | kinase | 264259, 264908, 264909, 264682, 22279000, 264488, 65274572, 22278995, 22278996, 22278997, 22278998, 22278999, 60432049, 264259, 29331822, 29331824, 29331825, 66714117, 29331826, 60432289, 35696052, 264905, 264509, 29331830, 265006, 60170831, 60432229, 60433356, 87168474, 265017, 265018, 265019, 264448, 264369, 264288, 21906765, 21906766, 55811957, 35695917, 265020, 265022, 52644150, 33657023, 65274620, 33657109, 18108370, 18108378, 55810784, 35696423, 55811576, 264556, 264558, 18108385, 60432113, 264583, 264584, 264585, 264586, 264587, 264907, 264512, 265011, 264683 |
| 2343 | 8775448 (4685, 4686) | Novel Protein sim. GBank gi 4929741jgb AAD34131.1j AF15189 - (AF151894) CGI-136 protein [Homo sapiens] | | UNCLASSIFIED | |
| 2344 | 79953198 (4687, 4688) | Novel Protein sim. GBank gi 2506307j sp P13944 CA1C_CHICK - COLLAGEN ALPHA 1(XII) CHAIN PRECURSOR (FIBROCHIMERIN) | Contains protein domain (PF00092) - von Willebrand factor type A domain | UNCLASSIFIED | 264758 |
| 2345 | 94318789 (4689, 4690) | | | collagen | 264488, 264259, 66712502, 264759, 83373044, 264566 |

| | | | | | |
|------|-----------------------|--|---|---------------|--|
| 2346 | 94131820 (4691, 4692) | Novel Protein sim. GBank gi 1255411 (U53153) - one short region of weak similarity to S. cerevisiae protease A inhibitor 3 (SP:P01094) and another short region of weak similarity to S. cerevisiae glucose repression mediator protein (SP:P14922) [Caenorhabditis elegans] | Contains protein domain (PF00515) - TPR Domain | proteaseinhib | 35696286, 22278998, 264259, 35696052, 29331828, 33657402, 60433356, 33109954, 87188559, 264603, 265019, 18108351, 264681, 264685, 21906768, 265021, 33657109, 55811576, 35695855, 264637, 52644332, 264557, 83373044, 22279000, 22279002 |
| 2347 | 85330367 (4693, 4694) | | | | 22278997, 264511, 264683, 264684, 264768, 264687, 264688, 264691, 264692, 55811576 |
| 2348 | 95196133 (4695, 4696) | Novel Protein sim. GBank gi 1929056[emb]CAA72805] - (Y12080) putative 3,4-dihydroxy-2-butanone kinase [Lycopersicon esculentum] | | kinase | 18108394, 35696286, 264259, 35696052, 264508, 264509, 264905, 264906, 264907, 264908, 264909, 264510, 264511, 265006, 265007, 264512, 265008, 265009, 264910, 264591, 264592, 264593, 264594, 264757, 264595, 264596, 264758, 265011, 264601, 264762, 18108351, 264764, 264288, 264766, 264768, 264689, 35695917, 264693, 264628, 18108370, 264628, 18108374, 35696423, 264631, 264635, 264636, 264637, 264638, 264639, 83373044, 18108385, 264567, 264486 |
| 2349 | 87776502 (4697, 4698) | Novel Protein sim. GBank gi 4884106[emb]CAB43254.1] - (AL050062) hypothetical protein [Homo sapiens] | | | 35696052, 29146499, 264908, 264369 |
| 2350 | 88260594 (4699, 4700) | | | | 22278998, 22278999, 264259, 29331822, 29331824, 29331825, 29331827, 29331828, 33109954, 21906754, 265010, 87188559, 265018, 265019, 284761, 264681, 264288, 18108357, 21906766, 21906767, 264691, 264692, 35695855, 87188518, 22279000, 22279002, 264482 |
| 2351 | 86968042 (4701, 4702) | Novel Protein sim. GBank gi 728832[sp]P39189[ALU2_HUMAN - III] ALU SUBFAMILY SB WARNING ENTRY III | | kinase | 56182575, 264809, 265006, 264558 |
| 2352 | 87337186 (4703, 4704) | Novel Protein sim. GBank gi 731637[sp]P38760[YHH5_YEAST - HYPOTHETICAL 75.9 KD PROTEIN IN SPO13-ARG4 INTERGENIC REGION] | Contains protein domain (PF00076) - RNA recognition motif. (a.k.a. RRM, RBD, or RNP domain) | UNCLASSIFIED | 264259, 264448 |
| 2353 | 91638784 (4705, 4706) | Novel Protein sim. GBank gi 1346855[sp]P48809[R827_DROME - HETEROGENEOUS NUCLEAR RIBONUCLEOPROTEIN 27C (HNRNP 48) (HRP48.1)] | Contains protein domain (PF00076) - RNA recognition motif. (a.k.a. RRM, RBD, or RNP domain) | dna_rna_bind | 29331826, 55812038, 265019, 264692, 264636 |
| 2354 | 87337189 (4707, 4708) | Novel Protein sim. GBank gi 731637[sp]P38760[YHH5_YEAST - HYPOTHETICAL 75.9 KD PROTEIN IN SPO13-ARG4 INTERGENIC REGION] | Contains protein domain (PF00076) - RNA recognition motif. (a.k.a. RRM, RBD, or RNP domain) | UNCLASSIFIED | 29331824, 264908, 265006, 265008 |

| | | | | | |
|------|-----------------------|--|--|--------------|--|
| 2355 | 91638786 (4709, 4710) | Novel Protein sim. GBank gjl4938503[emb]CAB43861.1] - (AL078465) hRNP-like protein [Arabidopsis thaliana] | Contains protein domain (PF000076) - RNA recognition motif. (a.k.a. RRM, RBD, or RNP domain) | dna_rna_bind | 56894075, 22278996, 35696286, 22278999, 264259, 29331825, 29331826, 29331828, 29146498, 264905, 264908, 265008, 264758, 87169474, 265010, 265017, 264687, 21908765, 21908767, 21908769, 264691, 264692, 263987, 18108370, 87168518, 22279000 |
| 2356 | 95327688 (4711, 4712) | Novel Protein sim. GBank gjl513920[gb]AAD40377.1] - (AF092135) PTD014 [Homo sapiens] | | | 52644507, 22278995, 35696286, 22278996, 22278997, 22278998, 22278999, 264259, 29331824, 66714117, 29331825, 60432289, 35696052, 29331828, 264908, 66712502, 264512, 265007, 265008, 60170831, 60432229, 60433356, 60433438, 264758, 52646317, 33109954, 21906754, 55811386, 87169474, 265017, 265018, 264605, 265019, 264681, 264682, 264448, 264369, 264288, 264686, 264768, 21906765, 21906766, 21906767, 21906768, 21906769, 265021, 60170815, 33857109, 27486264, 35695763, 55810764, 18108379, 35696423, 55811578, 35695855, 60170394, 56182323, 83373044, 18108385, 56526486, 264404, 60432113, 22279000, 264482, 264563, 264566, 264486, 264567 |
| 2357 | 87775458 (4713, 4714) | Novel Protein sim. GBank gjl4928741[gb]AAD34131.1]AF15189 - (AF151894) CGI-136 protein [Homo sapiens] | | UNCLASSIFIED | 264488, 264769, 18108394, 264259, 29331822, 18108370, 18108374, 264510, 265017, 264482, 264563, 264762, 264565, 264566, 264359, 18108354 |
| 2358 | 87777078 (4715, 4716) | Novel Protein sim. GBank gjl4218005 (AC006135) - putative vicilin storage protein (globulin-like) [Arabidopsis thaliana] | | UNCLASSIFIED | 22278997, 22278999, 264509, 264905, 264592, 18108351, 264681, 264682, 264769, 32833986, 18108374, 264556, 18108385, 264482 |
| 2359 | 87755859 (4717, 4718) | Novel Protein sim. GBank gjl1086830 (U41264) - coded for by C. elegans cDNA yk2018.5; coded for by C. elegans cDNA yk44g1.5; coded for by C. elegans cDNA yk12b7.5; coded for by C. elegans cDNA yk36g6.5; coded for by C. elegans cDNA yk2018.5; coded for by C. elegans cDNA yk16g12.... | | UNCLASSIFIED | 35696286, 22278998, 264905, 264511, 265007, 265008, 60433438, 264288, 264686, 21908769, 265020, 264692, 35695855, 264558, 56526486, 264563 |
| 2360 | 80046125 (4719, 4720) | Novel Protein sim. GBank gjl3881545[emb]CAA93779] - (Z59804) cDNA EST yk428d5.3 comes from this gene; cDNA EST yk428d5.5 comes from this gene [Caenorhabditis elegans] | | UNCLASSIFIED | 22278997, 29331826, 263981, 22279000 |
| 2361 | 84232191 (4721, 4722) | Novel Protein sim. GBank gjl746487 (U23514) - No definition line found [Caenorhabditis elegans] | | | 22278995, 22278999, 264512, 265009, 264757, 21906765, 65274620, 18108370, 60431528, 18108374, 264635, 60170394, 264482 |

| | | | | | |
|------|-----------------------|--|--|--------------|---|
| 2362 | 91721193 (4723, 4724) | Novel Protein sim. GBank gil1171083[sp]P19706[MYSB.ACACA - MYOSIN HEAVY CHAIN IB (MYOSIN HEAVY CHAIN IL) | | UNCLASSIFIED | 22278998, 264259, 29331822, 29331824, 60432289, 264509, 264512, 60432229, 60433356, 264448, 264682, 264683, 264369, 21906765, 21906768, 21906769, 60432113, 22279000, 22279002 |
| 2363 | 95006635 (4725, 4726) | Novel Protein sim. GBank (X83413) U88 [Human herpesvirus 6] | | UNCLASSIFIED | 264907, 264629, 264635 |
| 2364 | 94827104 (4727, 4728) | Novel Protein sim. GBank gil5639830[gb]AAD45886.1[AF14601 - (AF146018) hydroxypyruvate reductase (Homo sapiens)] | Contains protein domain (PF00338) - D-isomer specific 2-hydroxyacid dehydrogenases | reductase | 264488, 18108394, 264887, 18108398, 22278998, 56994075, 3569286, 22278997, 22278998, 264259, 66714117, 29331825, 35698052, 264509, 264905, 264906, 264907, 264908, 66712502, 264909, 264511, 265006, 264512, 265007, 265008, 33657402, 264758, 21906754, 87168474, 265010, 87168559, 264603, 265017, 265018, 265019, 264760, 284762, 18108351, 284448, 264764, 264683, 264684, 264288, 18108355, 264766, 18108358, 264689, 18108359, 21906765, 21906766, 21906767, 35695917, 265020, 265021, 265022, 60170615, 52644150, 264691, 33657023, 264692, 18108364, 33657109, 18108368, 18108370, 18108374, 35698423, 35695855, 264635, 264556, 264557, 264639, 60170394, 83373044, 18108383, 18108384, 18108385, 18108388, 56526486, 264482, 264564, 264486 |
| 2365 | 94140746 (4729, 4730) | Novel Protein sim. GBank gil1840045 (U49082) - transporter protein [Homo sapiens] | | transport | 22278998, 22278998, 22278999, 264907, 264909, 264910, 33657402, 264758, 264600, 264766, 264687, 264689, 21906765, 21906767, 21906768, 21906769, 265021, 33657023, 33657109, 83373044, 264566 |
| 2366 | 94312388 (4731, 4732) | | | UNCLASSIFIED | 52644507, 264259, 29331824, 29331825, 29331826, 29331827, 29331828, 264907, 29331830, 264909, 264511, 265008, 33657402, 264595, 52646317, 265017, 265018, 265019, 264605, 264685, 264766, 264689, 21906766, 21906769, 35695917, 265020, 265021, 265022, 52644150, 35695855, 52644332, 18108385, 18108387, 264564, 264568 |
| 2367 | 94140910 (4733, 4734) | Novel Protein sim. GBank gil1065457 (U40410) - C54G7.4 gene product [Caenorhabditis elegans] | Contains protein domain (PF00400) - WD domain, G-beta repeat | UNCLASSIFIED | 35698286, 21908768, 55810764, 65274791, 264567 |
| 2368 | 94322190 (4735, 4736) | | | | 264628 |

| | | | | |
|------|-----------------------|---|---|--|
| 2369 | 94314334 (4737, 4738) | Novel Protein sim. GBank gjl5360901[dbj BAA82158.1] - (AB029343) a-helix coiled-coil rod homologue [Homo sapiens] | struct | 52644507, 52646842, 35696286, 264092, 264094, 52645080, 35696052, 264107, 29331830, 52644045, 265006, 265007, 265009, 52644296, 52644229, 264689, 21906765, 21906768, 35695917, 265020, 52644150, 263967, 33657109, 27486265, 35695763, 18108370, 263974, 18108374, 18108376, 52644332, 263981, 18108385, 264508, 264909, 264596, 264369 |
| 2370 | 79804120 (4739, 4740) | | UNCLASSIFIED | |
| 2371 | 57280406 (4741, 4742) | | UNCLASSIFIED | 263967, 263981 |
| 2372 | 87642413 (4743, 4744) | | UNCLASSIFIED | 29331826, 265010, 265019, 35695917, 264634, 60432113 |
| 2373 | 87418611 (4745, 4746) | Novel Protein sim. GBank gjl458582[dbj BAA76813.1] - (AB023186) KIAA0969 protein [Homo sapiens] | | |
| 2374 | 94123665 (4747, 4748) | Novel Protein sim. GBank gjl5105131[dbj BAA80445.1] - (AF000061) 246aa long hypothetical ribonuclease PH [Aeropyrum pernix] | Contains protein domain (PF01138) - 3' exoribonuclease family | 264766, 35695917, 35695855, 263981, 264557, 264565 |
| 2375 | 87731355 (4749, 4750) | Novel Protein sim. GBank gjl1351115[sp P47758]SRPB_MOUSE - SIGNAL RECOGNITION PARTICLE RECEPTOR BETA SUBUNIT (SR-BETA) | UNCLASSIFIED | 60432049, 29331824, 264807, 52644045, 264512, 60433356, 21906754, 52644296, 87168559, 264448, 21906765, 21906768, 21906769, 33657023, 18108368, 55811576, 52644332 |
| 2376 | 87613744 (4751, 4752) | Novel Protein sim. GBank gjl2645435 (AF007780) - CHD3 [Drosophila melanogaster] | Contains protein domain (PF00628) - ATase-associated PHD-finger | 264259, 29331830, 264909, 264910, 265009, 60433438, 21906754, 265017, 265018, 265019, 264682, 264288, 264685, 21906767, 263972, 35695855, 87168518, 60432113 |
| 2377 | 95319689 (4753, 4754) | Novel Protein sim. GBank gjl5257005[gb AAD41239.1] - (AF083249) Rb binding protein homolog [Homo sapiens] | UNCLASSIFIED | 264095, 29331822, 29147620, 29331824, 66714117, 29331825, 29331826, 29331828, 33656970, 29146498, 29146499, 264509, 265006, 265007, 265008, 265009, 60170831, 265010, 265011, 265018, 55811150, 18108351, 264764, 264288, 21906767, 21906768, 29148627, 29148629, 265021, 33657023, 33657109, 18108370, 18108374, 18108379, 35696423, 264556, 83373044, 18108385, 18108388, 56526486, 22279000, 22279002, 264563 |
| 2378 | 94137032 (4755, 4756) | Novel Protein sim. GBank gjl1072198 (U40942) - No definition line found [Caenorhabditis elegans] | UNCLASSIFIED | 65274572, 56182575, 35696286, 264259, 29331822, 29331824, 66714117, 29331825, 29331826, 60432289, 29331827, 56182435, 264510, 265009, 60433356, 87168474, 265011, 265018, 264288, 21906765, 33657023, 264557, 56182323, 83373044, 18108385, 22279002, 264482 |
| 2379 | 65444324 (4757, 4758) | Novel Protein sim. GBank gjl3337357 (AC004481) - hypothetical protein [Arabidopsis thaliana] | Contains protein domain (PF00400) - kinase WD domain, G-beta repeat | 265017, 264288, 21906768 |

| | | | | | |
|------|-----------------------|--|--|--------------|--|
| 2380 | 86923062 (4759, 4760) | Novel Protein sim. GBank gij4502839[ref]NP_001845.1pCOL1 - collagen, type XI, alpha 1 | Contains protein domain (PF01410) - Fibrillar collagen C-terminal domain | collagen | 264908, 264910, 265011 |
| 2381 | 87608241 (4761, 4762) | Novel Protein sim. GBank gij445609[emb]CAB36555) - (AL031846) dJ742C19.5 (novel Chromobox protein) [Homo sapiens] | Contains protein domain (PF00385) - 'chromo' (CHR)romatin Organization M(oder)ifier domain | helicase | 56182575, 264091, 264093, 264259, 29331825, 264105, 264806, 60433356, 21906754, 265017, 265019, 264683, 264289, 264685, 264686, 264687, 264691, 264692, 264693, 55811576, 264636, 264567 |
| 2382 | 91225982 (4763, 4764) | Novel Protein sim. GBank gij4325130[gb]AAD17276j - (AF119716) dMi-2 protein [Drosophila melanogaster] | Contains protein domain (PF00628) - PHD-finger | transport | 29331824, 60432289, 264905, 264596, 21906754, 264769, 265022, 264693, 263967, 33657109, 264629, 264631, 264556, 83373044, 60432113, 264482 |
| 2383 | 87442841 (4765, 4766) | Novel Protein sim. GBank gij1902982[dbj]BAA19005) - (D89049) lecln-like oxidized LDL receptor [Bos taurus] | Contains protein domain (PF00059) - Lectin C-type domain | glycoprotein | 265009, 21906765, 21906766 |
| 2384 | 95354766 (4767, 4768) | Novel Protein sim. GBank gij2462851 (AF016252) - Spinophilin [Rattus norvegicus] | Contains protein domain (PF00595) - PDZ domain (Also known as DHR or GLGF) | struct | 264488, 52644507, 52645156, 52646365, 35696286, 22278999, 52645080, 29331824, 29331826, 35696052, 29331828, 264906, 264828, 52644045, 265006, 265008, 265009, 33109954, 33657084, 52644296, 265011, 265017, 265018, 264683, 52644229, 21906765, 21906767, 21906768, 265020, 52644150, 33657023, 264693, 65274620, 52645129, 33657109, 33657182, 27486261, 27486262, 27486264, 33657349, 27486265, 35695763, 18108374, 35695855, 264634, 264555, 264556, 264557, 52644332, 264558, 264559, 18108385, 22279000, 22279002 |
| 2385 | 95419485 (4769, 4770) | | UNCLASSIFIED | | 264488, 52644507, 52645156, 264887, 52646365, 22278995, 22278996, 22278997, 22278999, 264259, 52645080, 29331822, 29331826, 35696052, 52644045, 265006, 265007, 265008, 265009, 264910, 60432229, 60433356, 52646317, 21906754, 265019, 264448, 264683, 264686, 264687, 264689, 21906765, 21906766, 21906767, 21906769, 55811957, 265021, 265022, 264690, 264691, 264692, 65274620, 33657109, 18108370, 264631, 52644332, 22279000, 22279002, 264563, 264565, 264567 |
| 2386 | 94742649 (4771, 4772) | Novel Protein sim. GBank gij4929699[gb]AAD34110.1[AF15187] CGI-115 protein [Homo sapiens] | | glycoprotein | 264488, 22278995, 22278996, 22278997, 264259, 29148498, 264112, 264511, 60170831, 60432229, 264595, 60433438, 87168474, 87168559, 264682, 21906765, 21906766, 21906767, 21906769, 29148629, 35695917, 265021, 264690, 33657109, 264628, 18108376, 83373044, 60432113, 22279000, 264564, 264566, 264487 |

| | | | | | |
|------|-----------------------|--|--|--------------------------------|---|
| 2387 | 14997980 (4773, 4774) | | | UNCLASSIFIED | 264634 |
| 2388 | 11424804 (4775, 4776) | | | UNCLASSIFIED | 264595 |
| 2389 | 95310850 (4777, 4778) | Novel Protein sim. GBank gil4758058[ref]NP_004372.1 pCREB - cAMP responsive element binding protein-like 1 | Contains protein domain (PF00170) - bZIP transcription factor | UNCLASSIFIED - dna_rna_bind | 264488, 22278998, 22278999, 264509, 264905, 264906, 264907, 264908, 264909, 265006, 264511, 264512, 264910, 264591, 21906754, 264601, 264604, 264761, 18108351, 264764, 264288, 264766, 264768, 264769, 21906765, 21906768, 264692, 264693, 35696423, 264635, 264636, 264555, 83373044, 22279000, 264488 |
| 2390 | 94320912 (4779, 4780) | Novel Protein sim. GBank gil1644239[dbj]BAA122231 - (D84103) mitochondrial DNA polymerase gamma [Homo sapiens] | Contains protein domain (PF00476) - DNA polymerase family A | UNCLASSIFIED - polymerase | 52644507, 56182575, 22278995, 35696286, 22278998, 22278997, 22278999, 29331822, 29331825, 29331826, 35696052, 264805, 52644045, 265009, 264758, 264759, 33109954, 52644296, 85658542, 265011, 265017, 265018, 264605, 52644229, 21906765, 21906767, 21906768, 21906769, 35695917, 52644150, 33657023, 33657109, 33657349, 35695763, 18108370, 18108374, 18108376, 35696423, 35695855, 264555, 52644332, 56182233, 60170394, 83373044, 56526488 |
| 2391 | 80036194 (4781, 4782) | | | UNCLASSIFIED | 263976 |
| 2392 | 94245016 (4783, 4784) | Novel Protein sim. GBank gil4240169[dbj]BAA74863.11 - (AB020647) KIAA0840 protein [Homo sapiens] | Contains protein domain (PF00560) - Leucine Rich Repeat | UNCLASSIFIED - nuclease | 35696286, 35696052, 264508, 264905, 264509, 264906, 264907, 264908, 264909, 264510, 264511, 264512, 264910, 265009, 264591, 264758, 264600, 264604, 264762, 264448, 264764, 264369, 264766, 264768, 264769, 264689, 35695917, 264629, 18108374, 263978, 35696423, 35695855, 264631, 264634, 264635, 264636, 264637, 264638, 60170394, 264639, 264565, 264486 |

| | | | | | |
|------|-----------------------|--|---|---------------|---|
| 2383 | 95302633 (4785, 4786) | Novel Protein sim. GBank gil4506667/jetlNP_000983.1pRPLP - ribosomal protein, large, P0 | Contains protein domain (PF00466) - Ribosomal protein L10 | ribosomalprot | 18108392, 60424179, 264489, 18108394, 18108397, 22278995, 56994075, 35696286, 22278996, 22278997, 22278999, 264093, 60432049, 264259, 29331822, 29147620, 20281099, 29331824, 29331825, 68714117, 60432289, 29331826, 29331827, 29331828, 35696052, 29146499, 264508, 264509, 264905, 264907, 264908, 66712502, 52644045, 264828, 264909, 56182435, 264112, 264113, 264510, 265006, 264511, 265007, 265008, 265009, 264910, 264591, 264593, 60433356, 264595, 60433438, 52646317, 33109954, 21906754, 55811386, 265010, 265011, 265017, 285018, 265019, 264681, 264762, 18108351, 264763, 264682, 264764, 264683, 264389, 264288, 18108354, 264766, 264686, 264687, 264688, 264689, 18108359, 21906765, 21906766, 21906767, 21906768, 21906769, 55811957, 29148629, 29148784, 35695917, 265021, 265022, 33657023, 264692, 264693, 18108364, 33657109, 18108368, 27486261, 27486262, 33657349, 35695763, 18108370, 263972, 264629, 18108374, 263977, 18108376, 263978, 55810764, 35696423, 35695855, 264634, 60431850, 264555, 264637, 264557, 263981, 264558, 18108381, 60170394, 35696286, 22278997, 22278998, 56182181, 35696052, 265006, 264592, 55811386, 265010, 265011, 265017, 265019, 264448, 264683, 264288, 21906765, 21906768, 21906769, 55811957, 35695917, 33657023, 65274620, 33657182, 33657349, 35695763, 18108374, 18108376, 55810764, 55811576, 35696423, 60170394, 18108385, 264584, 264566, 264567 |
| 2394 | 94323266 (4787, 4788) | Novel Protein sim. GBank gil4159888 (AC004908) - zinc finger protein from gene of uncertain exon structure; similar to Q99676 (P1D:93025333) [Homo sapiens] | Contains protein domain (PF00096) - Zinc finger, C2H2 type | dna_ma_bind | 264259, 29331824, 264910, 264288, 265021, 83373044, 18108387, 264563, 264566 |
| 2395 | 95287212 (4789, 4790) | Novel Protein sim. GBank gil57127561/AAD47636.1(AF160798) calcium transporter CaT1 [Rattus norvegicus] | | dna_ma_bind | |

| | | | | | |
|------|-----------------------|---|--|--------------|---|
| 2398 | 95086700 (4791, 4792) | Novel Protein sim. GBank gij105322(pirj)B34087 - hypothetical protein (L1H 3' region) - human | Contains protein domain (PF00560) - Leucine Rich Repeat | nuclease | 52646365, 18108397, 56182575, 35698286, 22278997, 22278999, 60432049, 264259, 29331822, 29331824, 29331825, 29331826, 29331827, 29331828, 264106, 264907, 29331830, 66712502, 264110, 60170831, 264591, 33657402, 60433438, 55812038, 33109954, 21908754, 33657084, 87168474, 265017, 265018, 265019, 264760, 264448, 264288, 264768, 52644229, 21908768, 21908767, 265020, 265021, 60170615, 264692, 33657023, 65274620, 52645129, 33657182, 27486262, 27486264, 27486265, 264629, 18108374, 35698423, 35698855, 264631, 264556, 52644332, 264558, 83373044, 18108388, 87168518, 22278002, 264482 |
| 2397 | 87280854 (4793, 4794) | | | | 52644507, 52645156, 56182575, 264259, 29147620, 264905, 264907, 264908, 264809, 264910, 264758, 52644296, 264603, 264604, 264762, 264681, 264764, 18108357, 264769, 21908768, 264693, 264628, 264635, 264638, 264639, 264564 |
| 2398 | 88047689 (4795, 4796) | Novel Protein sim. GBank gij3258609 (AC005178) - H53_GS1 [Homo sapiens] | | UNCLASSIFIED | |
| 2399 | 87738965 (4797, 4798) | Novel Protein sim. GBank gij786117 (L41834) - nuclear protein [Ensis minor] | | UNCLASSIFIED | 56994075, 22278997, 22278998, 22278999, 264259, 29331822, 29331824, 66714117, 60432289, 264906, 29331830, 56182435, 264112, 264910, 33109954, 21908754, 87168474, 264600, 265017, 265018, 265019, 264764, 264765, 21908765, 21908768, 21908767, 21908769, 35695917, 265020, 265022, 60170815, 33657023, 18108370, 18108374, 264556, 60170394, 264558, 87168518, 22279000, 22279002, 264564, 264566, 264487 |
| 2400 | 91214116 (4799, 4800) | Novel Protein sim. GBank gij2352822(pirj)AAB69285.1) - (AF008945) glucose-6-phosphatase [Haplochromis nubilus] | | phosphatase | 21908766, 52646842, 56994075, 33657182, 27486262, 52644296, 265017 |

| | | | | | |
|------|-----------------------|--|---|---------------|--|
| 2401 | 91214118 (4801, 4802) | Novel Protein sim. GBank gij2352822[pb]AAB69285, 1] - (AF008845) glucose-6-phosphatase [Haplochromis nubilus] | | phosphatase | 52644507, 52645156, 52644229, 264688, 21906764, 21906765, 52646385, 52646842, 21906766, 21906767, 21906768, 22278995, 35695917, 56994075, 35696286, 22278996, 22278997, 265020, 22278998, 22278999, 264259, 33657023, 52645080, 264693, 29331824, 33637109, 52645129, 29331826, 33657182, 29331827, 35696052, 27486261, 27486262, 33656970, 33657349, 27486285, 35695763, 264106, 264905, 35696423, 35695855, 265006, 265007, 265008, 265009, 264637, 52644332, 56812038, 52646317, 18108385, 52644296, 87168474, 265010, 87168559, 80432113, 265017, 265018, 265019, 264563, 264288, 264907, 264908, 264909, 264566 |
| 2402 | 91221408 (4803, 4804) | Novel Protein sim. GBank gij4689259[pb]AAD27832, 1[AF12185 - (AF121859) sorting nexin 9 [Homo sapiens] | | | |
| 2403 | 94135432 (4805, 4806) | Novel Protein sim. GBank gij4929575[pb]AAD34048, 1[AF15181 - (AF151811) CGI-53 protein [Homo sapiens] | Contains protein domain (PF000062) - C-type lysozyme/alpha-lactalbumin family | | 22278999, 35696052, 265018, 264686, 264693, 83373044, 264567 |
| 2404 | 95312605 (4807, 4808) | Novel Protein sim. GBank gij2315786 (AF016685) - similar to short chain-type dehydrogenases [Caenorhabditis elegans] | Contains protein domain (PF00106) - short chain dehydrogenase | dehydrogenase | 35696286, 29331826, 35696052, 265008, 265018, 21908769, 264584 |
| 2405 | 94311851 (4809, 4810) | Novel Protein sim. GBank gij464178[d]BAA03581] - (D14853) polyprotein [Hepatitis C virus] | | UNCLASSIFIED | 35696286, 29331822, 265007, 21906754, 265017, 265018, 265019, 264763, 264369, 21906765, 35695917, 265020, 265021, 52644150, 264693, 35695855, 264632, 52644332, 22279002 |
| 2406 | 88094501 (4811, 4812) | Novel Protein sim. GBank gij2773363 (AF041382) - microtubule binding protein D-CLIP-190 [Drosophila melanogaster] | Contains protein domain (PF01302) - CAP-Gly domain | struct | 52646842, 22278994, 22278996, 35696286, 22278997, 22278998, 22278999, 264092, 264093, 60432049, 264259, 29331822, 29331824, 29331825, 29331827, 35696052, 29331828, 264102, 264106, 264906, 52644045, 265007, 265008, 265009, 264910, 264592, 60433356, 6043438, 33109954, 265010, 265011, 265018, 265019, 264369, 264685, 264686, 21906768, 21906769, 52644150, 264693, 52645129, 264628, 35696423, 264632, 56182323, 264639, 22279000, 22279002, 264563, 264685, 264686 |
| 2407 | 79465005 (4813, 4814) | | | UNCLASSIFIED | |
| 2408 | 67351503 (4815, 4816) | Novel Protein sim. GBank gij423442[p]S33513 - gene Fil protein - mouse | | UNCLASSIFIED | 264910, 265010, 264448, 264557 |

| | | | | | |
|------|-----------------------|---|--|--------------|--|
| 2409 | 94741770 (4817, 4818) | Novel Protein sim. GBank gij1178801spj45966jYNZ6 CAEL - HYPOTHETICAL 20.8 KD PROTEIN T09A5.6 IN CHROMOSOME III | | UNCLASSIFIED | 22278995, 22278996, 22278997, 264097, 29331822, 29331824, 29331827, 29148498, 52644045, 60433438, 33657084, 87168474, 264760, 21906767, 29148627, 29148629, 52644150, 33657023, 263967, 20281069, 18108374, 20281071, 56182323, 83373044, 18108385, 87168518 22278998, 264259 |
| 2410 | 87604860 (4819, 4820) | Novel Protein sim. GBank gij4966262jgbAAC48052.21 - (U84849) Contains similarity to Pfam domain: PF00846 (F- box), Score=28.7, E-value=4.3e-05, N=1 [Caenorhabditis elegans] | Contains protein domain (PF00080) - Copper/zinc superoxide dismutase (SODC) | | |
| 2411 | 87534633 (4821, 4822) | Novel Protein sim. GBank gij3114713 (AF061346) - Edp1 protein [Mus musculus] | | Inf | 29331824, 29331827, 29331828, 264764, 264369, 33657109, 56182323 |
| 2412 | 87778332 (4823, 4824) | Novel Protein sim. GBank gij5410336jgbAAD43038.1) - (AF106685) myelin gene expression factor 2 [Homo sapiens] | Contains protein domain (PF00076) - RNA recognition motif (a.k.a. RRM, RBD, or RNP domain) | dna_rna_bind | 22278998, 29331827, 264907, 265011, 265017, 265018, 265019, 18108351, 21906766, 265020, 33657109, 264559, 18108385 |
| 2413 | 94133820 (4825, 4826) | Novel Protein sim. GBank gij5262705jemb[CAB45778.1] - (AL080214) hypothetical protein [Homo sapiens] | Contains protein domain (PF00038) - Intermediate filament proteins | struct | 264488, 264259, 29331826, 264508, 264905, 264509, 264908, 264907, 264908, 264510, 264511, 264512, 265008, 265009, 264910, 265011, 264682, 264764, 264766, 264686, 264768, 264688, 265021, 33657023, 18108370, 264628, 35695855, 264632, 264634, 264635, 264636, 83373044, 264563, 264564, 264565, 264566, 264567, 264486 |
| 2414 | 94312590 (4827, 4828) | Novel Protein sim. GBank gij1082340jprj[S52863 - DNA- binding protein R kappa B - human | | ubiquitin | 52645156, 52646842, 52646365, 18108398, 56182575, 22278994, 22278995, 56994075, 22278998, 35696286, 22278997, 22278998, 22278999, 264259, 52645080, 29331822, 29331824, 29331825, 29331826, 29331827, 29331828, 35696052, 33656970, 29331830, 264909, 60433356, 33657402, 264594, 52646317, 21906754, 33657084, 265010, 87168559, 265017, 265018, 265019, 264369, 264684, 264687, 264688, 56181562, 21906764, 264689, 21906765, 21906766, 21906767, 29148627, 21906769, 265020, 265021, 60170615, 33657023, 264693, 52645129, 33657109, 33657182, 27486261, 27486262, 27486265, 33657349, 18108370, 60431528, 264629, 18108374, 18108378, 55810764, 264636, 52644332, 264638, 264558, 56182323, 83373044, 18108385, 87168518, 22279002 |

| | | | | | |
|------|-----------------------|--|--|--|--|
| 2415 | 88086002 (4829, 4830) | Novel Protein sim. GBank gij423915[pil]A45439 - myosin I heavy chain - rat | Contains protein domain (PF00063) - struct | Myosin head (motor domain) | 264259, 264908, 60433356, 33657402, 21908754, 265018, 264687, 264689, 21908769, 55811957, 265021, 264690, 264691, 33657023, 264693, 35698423, 56182323, 56526488 |
| 2416 | 94118358 (4831, 4832) | Novel Protein sim. GBank gij3025445 (AC004528) - R32184_1 [Homo sapiens] | | | 264638 |
| 2417 | 87733334 (4833, 4834) | Novel Protein sim. GBank gij1084944[pil]S54495 - hypothetical protein YPR021c - yeast (Saccharomyces cerevisiae) | Contains protein domain (PF00153) - | Contains protein domain (PF00153) - Mitochondrial carrier proteins | 264094, 29331822, 29331824, 29331827, 264369 |
| 2418 | 94234349 (4835, 4836) | Novel Protein sim. GBank gij1176572[sp]P45895[YNA4, CAEEL - HYPOTHETICAL 91.0 KD PROTEIN PAR.2 IN CHROMOSOME III | UNCLASSIFIED | Contains protein domain (PF00411) - Ribosomal protein S11 | 56994075, 264091, 264259, 29331824, 29331825, 60432289, 29331828, 264905, 264907, 264511, 265009, 60432229, 21908754, 87168559, 265019, 264682, 21908768, 21908769, 265020, 265021, 33657023, 65274620, 18108370, 55811576, 264634, 60170394, 18108385, 22279000, 22279002, 264568 |
| 2419 | 82374249 (4837, 4838) | Novel Protein sim. GBank gij284006[pil]S18732 - autoantigen, 64K - human | struct | | 264569, 264762, 264448, 264691, 264631, 264634, 264555, 264558, 264638, 264558 |
| 2420 | 94844244 (4839, 4840) | Novel Protein sim. GBank gij1076211[pil]S50755 - hypothetical protein VSP-3 - Chlamydomonas reinhardtii | UNCLASSIFIED | | 29331824, 29331825, 29331828, 60432229, 33109954, 85658542, 87168474, 265018, 264288, 265020, 264564 |
| 2421 | 87805345 (4841, 4842) | Novel Protein sim. GBank gij2224567[dbj]BAA20772 - (AB002311) KIAA0313 [Homo sapiens] | UNCLASSIFIED | | 264909, 264768, 264638 |
| 2422 | 88084714 (4843, 4844) | Novel Protein sim. GBank gij2224567[dbj]BAA20772 - (AB002311) KIAA0313 [Homo sapiens] | transport | Contains protein domain (PF00617) - RasGEF domain | 18108392, 18108394, 18108398, 264906, 265006, 265010, 18108351, 18108374, 18108385 |
| 2423 | 88058390 (4845, 4846) | Novel Protein sim. GBank gij4505153[ref]NP_002392.1pMEKK - MAP/ERK kinase kinase 3 | kinase | Contains protein domain (PF00069) - Eukaryotic protein kinase domain | 264259, 60432049, 29331822, 29331826, 60432289, 29331828, 265008, 265009, 60433356, 21908754, 265017, 265018, 265019, 21908766, 21908768, 21908769, 265020, 265021, 20281149, 263971, 60432113 |
| 2424 | 94854047 (4847, 4848) | Novel Protein sim. GBank gij2988398 (AC004381) - Unknown gene product [Homo sapiens] | UNCLASSIFIED | | 56182575, 35696286, 22278997, 60432049, 264259, 29331826, 29331828, 264905, 66712502, 29331830, 60433356, 265011, 265019, 264768, 21908768, 55811957, 264692, 33657023, 33657109, 55811576, 56182323, 83373044, 18108385, 18108388, 60432113, 22279000 |
| 2425 | 87415981 (4849, 4850) | Novel Protein sim. GBank gij2077932[dbj]BAA19879 - (D86556) Protein Kinase [Rattus norvegicus] | kinase | Contains protein domain (PF00069) - Eukaryotic protein kinase domain | 264634 |
| 2426 | 87613945 (4851, 4852) | Novel Protein sim. GBank gij2039368[gb]AA053003.1 - (U94619) circulating cathodic antigen [Schistosoma mansoni] | UNCLASSIFIED | | 22278996, 22278998, 264259, 264102, 264512, 265008, 21908767, 18108370, 18108374, 263976 |

| | | | | | |
|------|-----------------------|---|--|------------------|--|
| 2427 | 87622683 (4853, 4854) | Novel Protein sim. GBank gij4680695[gib]/AAD27737.1[AF13298 - (AF132982) CGI-28 protein [Homo sapiens] | Contains protein domain (PF00573) - Ribosomal protein L4/L1 family | ribosomalprot | 264259, 20281099, 35696052, 265008, 264594, 265011, 264760, 18108351, 264682, 264683, 284369, 264684, 264686, 264687, 264689, 21906766, 264691, 264692, 18108374, 18108377, 264557, 264639, 18108385 |
| 2428 | 85732889 (4855, 4856) | Novel Protein sim. GBank gij1537070 (U63840) - nucleoporin p54 [Rattus norvegicus] | | | 22278996, 22278999, 35696052, 21906754, 264288, 21906765, 21906788, 21906789, 35695917, 265020, 263972, 22279002 |
| 2429 | 87769276 (4857, 4858) | Novel Protein sim. GBank gij601931 (M94316) - neurofilament-H [Oryctolagus cuniculus] | Contains protein domain (PF00711) - Beta defensins | UNCLASSIFIED | 22278999, 29331824, 264806, 264909, 264511, 265009, 21906754, 265017, 265018, 265019, 264446, 264883, 264288, 21906765, 21906768, 265021, 264693, 18108381 |
| 2430 | 86948827 (4859, 4860) | | | UNCLASSIFIED | 264112, 264691 |
| 2431 | 87649884 (4861, 4862) | Novel Protein sim. GBank gij3860729[emb]CAA14630] - (AJ235270) CELL DIVISION PROTEIN FTSJ (ftsJ) [Rickettsia prowazekii] | Contains protein domain (PF01728) - FtsJ cell division protein | | 29331826, 29331827, 35696052, 29146499, 284905, 284906, 264681, 264288, 264689, 21906765, 264692, 35696423 |
| 2432 | 80083033 (4863, 4864) | Novel Protein sim. GBank gij3876367[emb]CAA93287] - (Z69360) Weak similarity to Eimeria thrombospondin (PIR Acc. No. A45517); cDNA EST EMBL:M89266 comes from this gene; cDNA EST yk295b9.5 comes from this gene [Caenorhabditis elegans] | | protease | 264634, 264558 |
| 2433 | 80055092 (4865, 4866) | Novel Protein sim. GBank gij2224593[dbj]BAA20784] - (AB002324) KIAA0326 [Homo sapiens] | Contains protein domain (PF00096) - Zinc finger, C2H2 type | transcriptfactor | 264589, 264905, 265018, 264762, 264683, 264691, 284556, 264557, 264639, 264558 |
| 2434 | 18520148 (4867, 4868) | | | UNCLASSIFIED | 264553 |
| 2435 | 20758044 (4869, 4870) | Novel Protein sim. GBank gij1263289 (U47856) - fibrin-4 [Araneus diadematus] | | UNCLASSIFIED | 265008, 264758, 265010, 264689, 27486261, 263972, 18108374, 18108381 |
| 2436 | 88044008 (4871, 4872) | Novel Protein sim. GBank gij3641352 (AF091234) - putative transcription factor [Mus musculus] | | UNCLASSIFIED | 29331828, 265007, 265009, 265017, 264760, 264685, 264693, 264585 |
| 2437 | 83363424 (4873, 4874) | | | UNCLASSIFIED | 29331826, 264508, 264905, 264509, 264906, 264907, 264908, 264909, 264511, 265006, 264512, 264910, 265009, 264591, 33657402, 21906754, 265011, 264760, 264764, 264685, 264686, 264768, 35695917, 33657023 |
| 2438 | 84143473 (4875, 4876) | Novel Protein sim. GBank gij3860014 (AF091088) - unknown [Homo sapiens] | Contains protein domain (PF01256) - Uncharacterized protein family UPF0031 | UNCLASSIFIED | 264693, 264631, 264632, 56182323, 264558, 83373044, 264563, 264564, 264565, 264566, 264567 |

| | | | | | |
|------|-----------------------|--|--|------------------|--|
| 2439 | 94850650 (4877, 4878) | Novel Protein sim. GBank gi 4263519 gb AAD15345 - (AC004044) small nuclear riboprotein Sm-D1 (Arabidopsis thaliana) | Contains protein domain (PF01423) - Sm protein | UNCLASSIFIED | 60424179, 18108397, 56182575, 56181686, 56994075, 22278996, 35696286, 22278997, 22278999, 264259, 52645080, 29331822, 56182181, 29331824, 60424269, 66714117, 29331825, 60432289, 29331826, 29331827, 29331828, 35696052, 29146498, 264509, 264905, 264906, 52644045, 60431735, 33109954, 21906754, 33657084, 55811386, 52644296, 87169474, 265017, 265018, 265019, 18108351, 264448, 264288, 264768, 52644229, 56181562, 21906764, 21906765, 21906766, 21906767, 21906768, 21906769, 35695917, 33657023, 33657109, 33657182, 27486262, 27486264, 33657349, 27486265, 35695763, 18108370, 60431528, 263977, 55810764, 35696423, 65274791, 35695855, 60431850, 56182323, 60432113, 22279000, 22279002, 264567 |
| 2440 | 87641733 (4879, 4880) | | | UNCLASSIFIED | 29331822, 29331824, 29331825, 29331827, 35696052, 264508, 264907, 264510, 265018, 265019, 264448, 264369, 265020, 265021, 56182323, 264639, 22279002 |
| 2441 | 87623914 (4881, 4882) | Novel Protein sim. GBank gi 3024889 sp P56524 Y288_HUMAN - HYPOTHETICAL PROTEIN KIAA0288 (HAG116) | | UNCLASSIFIED | 264488, 264629, 18108374, 264564 |
| 2442 | 87273550 (4883, 4884) | Novel Protein sim. GBank gi 4506013 ref NP_002703.1 pPPP1 - protein phosphatase 1, regulatory subunit 7 | Contains protein domain (PF00560) - Leucine Rich Repeat | UNCLASSIFIED | 22278996, 22278999, 29331822, 264768, 264693 |
| 2443 | 84305949 (4885, 4886) | Novel Protein sim. GBank gi 1170658 sp Q02975 KID1_RAT - RENAL TRANSCRIPTION FACTOR KID-1 (TRANSCRIPTION FACTOR 17) | Contains protein domain (PF01352) - KRAB box | transcriptfactor | 264906 |
| 2444 | 88086345 (4887, 4888) | Novel Protein sim. GBank gi 4758824 ref NP_004280.1 pNRF3 - nuclear factor (erythroid-derived 2)-like 3 | Contains protein domain (PF00170) - bZIP transcription factor | transcriptfactor | 264259, 18108382, 18108383, 18108385, 22279000 |
| 2445 | 87338636 (4889, 4890) | Novel Protein sim. GBank gi 2135950 pir I S58222 - PQ-rich protein - human | | | 264259, 35696052, 264369, 18108361 |
| 2446 | 88059293 (4891, 4892) | Novel Protein sim. GBank gi 4753887 emb CAA05409.2 - (AJ002424) p85 protein [Rattus norvegicus] | Contains protein domain (PF00095) - WAP-type (Whey Acidic Protein) 'four-disulfide core' | proteaseinhib | 265011, 264689, 33657023, 263981, 18108385 |
| 2447 | 94845149 (4893, 4894) | Novel Protein sim. GBank gi 4885613 ref NP_005409.1 pST5 - suppression of tumorigenicity 5 | | cadherin | 56182575, 264259, 29331824, 29331825, 29331827, 60433356, 60433438, 264758, 265018, 264692, 65274620, 60431528, 65274791, 56182323 |

| | | | | | |
|------|-----------------------|--|---|--------------|--|
| 2448 | 87749680 (4895, 4896) | | | UNCLASSIFIED | 22278996, 22278997, 22278999, 29331828, 35896052, 264107, 264110, 87168474, 87168559, 18108351, 21906767, 21906769, 27486282, 263976 |
| 2449 | 87869075 (4897, 4898) | Novel Protein sim. GBank gij28837[sp P39194 ALU7_HUMAN - IIII ALU SUBFAMILY SQ WARNING ENTRY IIII | | cadherin | 264259, 264828, 265007, 264595, 265021, 56526486 |
| 2450 | 86597784 (4899, 4900) | | | UNCLASSIFIED | 264906 |
| 2451 | 91014583 (4901, 4902) | Novel Protein sim. GBank gij1710021[sp P35290 RB24_MOUSE - RAS-RELATED PROTEIN RAB-24 (RAB-16) | Contains protein domain (PF00071) - Ras family | glycoprotein | 264093, 29331822, 29331824, 29331825, 66714117, 29331826, 29331828, 35896052, 264907, 66712502, 29331830, 264910, 265009, 284758, 265017, 265018, 264762, 264448, 264288, 21906767, 265021, 33657023, 264693, 33657109, 263969, 83373044, 18108385 |
| 2452 | 91230508 (4903, 4904) | Novel Protein sim. GBank gij1504034[dbj BAA13216 - (D86980) KIAA0227 [Homo sapiens] | | isomerase | 264102, 264112, 264688, 263972, 18108374, 83373044, 264583 |
| 2453 | 84201088 (4905, 4906) | Novel Protein sim. GBank gij2880079 (AC004142) - similar to murine leucine-rich repeat protein; possible role in neural development by protein-protein interactions; 93% similarity to D49802 (PID:g1369906) [Homo sapiens] | Contains protein domain (PF00560) - Leucine Rich Repeat | ngfirecep | 264509, 264512, 18108385 |
| 2454 | 95310691 (4907, 4908) | Novel Protein sim. GBank gij1076802[pir S49915 - extensin like protein - maize | Contains protein domain (PF00170) - bZIP transcription factor | UNCLASSIFIED | 263994, 66714117, 29331827, 264508, 264509, 284905, 264908, 284907, 264908, 264909, 264510, 264511, 264512, 265009, 264910, 264591, 264758, 264759, 265010, 265011, 264603, 264604, 264760, 264761, 264762, 18108351, 264764, 264765, 264766, 264686, 264768, 264769, 264534, 264691, 264692, 33657023, 264693, 33657109, 264628, 263978, 35895855, 264634, 264635, 264637, 264638, 264639, 83373044, 18108385, 264563, 264564, 264486 |
| 2455 | 95288301 (4909, 4910) | Novel Protein sim. GBank gij543817[sp P35585 AP47_MOUSE - CLATHRIN COAT ASSEMBLY PROTEIN AP47 (CLATHRIN COAT ASSOCIATED PROTEIN AP47) (GOLGI ADAPTOR AP-1 47 KD PROTEIN) (HA1 47 KD SUBUNIT) (CLATHRIN ASSEMBLY PROTEIN ASSEMBLY PROTEIN COMPLEX 1 MEDIUM CHAIN) | Contains protein domain (PF00828) - Adaptor complexes medium subunit family | glycoprotein | 264488, 22278996, 264259, 35896052, 264905, 264906, 264907, 264908, 264909, 264510, 264512, 265008, 265009, 264910, 264591, 264592, 264593, 264594, 264758, 265019, 264760, 264681, 18108351, 264683, 264764, 18108354, 264766, 264768, 264769, 264689, 21906766, 21906767, 21906769, 29148629, 35895917, 265020, 265022, 33657023, 33657109, 18108370, 264628, 264629, 264631, 264632, 264635, 56182323, 60170394, 18108385, 264563, 264564, 264565, 264567 |
| 2456 | 88166700 (4911, 4912) | Novel Protein sim. GBank gij2588630 (AC003079) - Ankyrin; like. 54% similar to 2022340A (NID:g1092123) in exons spanning 43974 to 11551 of clone. [Homo sapiens] | Contains protein domain (PF00023) - Ank repeat | kinase | 264693 |

| | | | | | |
|------|-----------------------|---|---|--------------|---|
| 2457 | 84118375 (4913, 4914) | Novel Protein sim. GBank gij3025447 (AC004528) - R32184_3 [Homo sapiens] | | UNCLASSIFIED | 56181686, 264905, 264907, 264511, 264596, 55811386, 264682, 264684, 264685, 264687, 264691, 33657023, 264693, 35695855, 264636, 264555, 56182323, 264558, 56526486, 264563 |
| 2458 | 85675304 (4915, 4918) | Novel Protein sim. GBank gij2384942 (AF022985) - Similar to collagen [Caenorhabditis elegans] | | UNCLASSIFIED | 264691, 264693, 264634, 264559 |
| 2459 | 87551913 (4917, 4918) | Novel Protein sim. GBank gij5441942igb/AAD43187.1 AC00499 - (AC004997) supported by mouse EST AA538043 (NID: g2284036) [Homo sapiens] | Contains protein domain (PF00097) - Zinc finger, C3HC4 type (RING finger) | transport | 27486265 |
| 2460 | 94315289 (4919, 4920) | Novel Protein sim. GBank gij4929701igb/AAD34111.1 AF15187 - (AF151874) CGI-116 protein [Homo sapiens] | | kinase | 55274572, 35696286, 22278996, 22278997, 60432049, 29331822, 29331824, 29331825, 66714117, 60432289, 29331826, 29331827, 29331828, 33656970, 29146499, 264102, 264109, 60433438, 265017, 265018, 265019, 264288, 21908765, 21908768, 21908769, 35695917, 265020, 264691, 33657023, 27486261, 18108374, 35695855, 87168518, 60432113 |
| 2461 | 87645147 (4921, 4922) | Novel Protein sim. GBank gij4426862igb/AAD20633 - (AF126082) Arf-like 2 binding protein BART1 [Homo sapiens] | | UNCLASSIFIED | 264259, 29331828, 264910, 18108351, 18108370, 18108374 |
| 2462 | 86998002 (4923, 4924) | Novel Protein sim. GBank gij5420387 emb CAB46679.1 - (AJ243459) proteophosphoglycan [Leishmania major] | | | 264809, 264758, 264684, 18108374, 264637, 18108385 |
| 2463 | 84388543 (4925, 4926) | Novel Protein sim. GBank gij5052516igb/AAD38588.1 AF14561 - (AF145613) BcDNA GH03108 [Drosophila melanogaster] | Contains protein domain (PF00096) - Zinc finger, C2H2 type | UNCLASSIFIED | 264681, 264566 |
| 2464 | 91219957 (4927, 4928) | Novel Protein sim. GBank gij5410300igb/AAD43021.1 - (AF100757) COP9 complex subunit 4 [Homo sapiens] | Contains protein domain (PF01399) - PCI domain | protease | 264489, 52646842, 22278995, 35696286, 22278996, 22278997, 22278999, 264259, 29331822, 29331824, 66714117, 29331825, 29331826, 29331827, 35696032, 29331828, 264509, 52644045, 264510, 264511, 264512, 265008, 60170831, 264593, 52646317, 33109954, 33657084, 265017, 265018, 265019, 264762, 264448, 264764, 264288, 264766, 21908765, 21908768, 21908769, 21908770, 21908771, 21908772, 21908773, 21908774, 21908775, 21908776, 21908777, 21908778, 21908779, 21908780, 21908781, 21908782, 21908783, 21908784, 21908785, 21908786, 21908787, 21908788, 21908789, 21908790, 21908791, 21908792, 21908793, 21908794, 21908795, 21908796, 21908797, 21908798, 21908799, 21908800, 21908801, 21908802, 21908803, 21908804, 21908805, 21908806, 21908807, 21908808, 21908809, 21908810, 21908811, 21908812, 21908813, 21908814, 21908815, 21908816, 21908817, 21908818, 21908819, 21908820, 21908821, 21908822, 21908823, 21908824, 21908825, 21908826, 21908827, 21908828, 21908829, 21908830, 21908831, 21908832, 21908833, 21908834, 21908835, 21908836, 21908837, 21908838, 21908839, 21908840, 21908841, 21908842, 21908843, 21908844, 21908845, 21908846, 21908847, 21908848, 21908849, 21908850, 21908851, 21908852, 21908853, 21908854, 21908855, 21908856, 21908857, 21908858, 21908859, 21908860, 21908861, 21908862, 21908863, 21908864, 21908865, 21908866, 21908867, 21908868, 21908869, 21908870, 21908871, 21908872, 21908873, 21908874, 21908875, 21908876, 21908877, 21908878, 21908879, 21908880, 21908881, 21908882, 21908883, 21908884, 21908885, 21908886, 21908887, 21908888, 21908889, 21908890, 21908891, 21908892, 21908893, 21908894, 21908895, 21908896, 21908897, 21908898, 21908899, 21908900, 21908901, 21908902, 21908903, 21908904, 21908905, 21908906, 21908907, 21908908, 21908909, 21908910, 21908911, 21908912, 21908913, 21908914, 21908915, 21908916, 21908917, 21908918, 21908919, 21908920, 21908921, 21908922, 21908923, 21908924, 21908925, 21908926, 21908927, 21908928, 21908929, 21908930, 21908931, 21908932, 21908933, 21908934, 21908935, 21908936, 21908937, 21908938, 21908939, 21908940, 21908941, 21908942, 21908943, 21908944, 21908945, 21908946, 21908947, 21908948, 21908949, 21908950, 21908951, 21908952, 21908953, 21908954, 21908955, 21908956, 21908957, 21908958, 21908959, 21908960, 21908961, 21908962, 21908963, 21908964, 21908965, 21908966, 21908967, 21908968, 21908969, 21908970, 21908971, 21908972, 21908973, 21908974, 21908975, 21908976, 21908977, 21908978, 21908979, 21908980, 21908981, 21908982, 21908983, 21908984, 21908985, 21908986, 21908987, 21908988, 21908989, 21908990, 21908991, 21908992, 21908993, 21908994, 21908995, 21908996, 21908997, 21908998, 21908999, 21909000, 21909001, 21909002, 21909003, 21909004, 21909005, 21909006, 21909007, 21909008, 21909009, 21909010, 21909011, 21909012, 21909013, 21909014, 21909015, 21909016, 21909017, 21909018, 21909019, 21909020, 21909021, 21909022, 21909023, 21909024, 21909025, 21909026, 21909027, 21909028, 21909029, 21909030, 21909031, 21909032, 21909033, 21909034, 21909035, 21909036, 21909037, 21909038, 21909039, 21909040, 21909041, 21909042, 21909043, 21909044, 21909045, 21909046, 21909047, 21909048, 21909049, 21909050, 21909051, 21909052, 21909053, 21909054, 21909055, 21909056, 21909057, 21909058, 21909059, 21909060, 21909061, 21909062, 21909063, 21909064, 21909065, 21909066, 21909067, 21909068, 21909069, 21909070, 21909071, 21909072, 21909073, 21909074, 21909075, 21909076, 21909077, 21909078, 21909079, 21909080, 21909081, 21909082, 21909083, 21909084, 21909085, 21909086, 21909087, 21909088, 21909089, 21909090, 21909091, 21909092, 21909093, 21909094, 21909095, 21909096, 21909097, 21909098, 21909099, 21909100, 21909101, 21909102, 21909103, 21909104, 21909105, 21909106, 21909107, 21909108, 21909109, 21909110, 21909111, 21909112, 21909113, 21909114, 21909115, 21909116, 21909117, 21909118, 21909119, 21909120, 21909121, 21909122, 21909123, 21909124, 21909125, 21909126, 21909127, 21909128, 21909129, 21909130, 21909131, 21909132, 21909133, 21909134, 21909135, 21909136, 21909137, 21909138, 21909139, 21909140, 21909141, 21909142, 21909143, 21909144, 21909145, 21909146, 21909147, 21909148, 21909149, 21909150, 21909151, 21909152, 21909153, 21909154, 21909155, 21909156, 21909157, 21909158, 21909159, 21909160, 21909161, 21909162, 21909163, 21909164, 21909165, 21909166, 21909167, 21909168, 21909169, 21909170, 21909171, 21909172, 21909173, 21909174, 21909175, 21909176, 21909177, 21909178, 21909179, 21909180, 21909181, 21909182, 21909183, 21909184, 21909185, 21909186, 21909187, 21909188, 21909189, 21909190, 21909191, 21909192, 21909193, 21909194, 21909195, 21909196, 21909197, 21909198, 21909199, 21909200, 21909201, 21909202, 21909203, 21909204, 21909205, 21909206, 21909207, 21909208, 21909209, 21909210, 21909211, 21909212, 21909213, 21909214, 21909215, 21909216, 21909217, 21909218, 21909219, 21909220, 21909221, 21909222, 21909223, 21909224, 21909225, 21909226, 21909227, 21909228, 21909229, 21909230, 21909231, 21909232, 21909233, 21909234, 21909235, 21909236, 21909237, 21909238, 21909239, 21909240, 21909241, 21909242, 21909243, 21909244, 21909245, 21909246, 21909247, 21909248, 21909249, 21909250, 21909251, 21909252, 21909253, 21909254, 21909255, 21909256, 21909257, 21909258, 21909259, 21909260, 21909261, 21909262, 21909263, 21909264, 21909265, 21909266, 21909267, 21909268, 21909269, 21909270, 21909271, 21909272, 21909273, 21909274, 21909275, 21909276, 21909277, 21909278, 21909279, 21909280, 21909281, 21909282, 21909283, 21909284, 21909285, 21909286, 21909287, 21909288, 21909289, 21909290, 21909291, 21909292, 21909293, 21909294, 21909295, 21909296, 21909297, 21909298, 21909299, 21909300, 21909301, 21909302, 21909303, 21909304, 21909305, 21909306, 21909307, 21909308, 21909309, 21909310, 21909311, 21909312, 21909313, 21909314, 21909315, 21909316, 21909317, 21909318, 21909319, 21909320, 21909321, 21909322, 21909323, 21909324, 21909325, 21909326, 21909327, 21909328, 21909329, 21909330, 21909331, 21909332, 21909333, 21909334, 21909335, 21909336, 21909337, 21909338, 21909339, 21909340, 21909341, 21909342, 21909343, 21909344, 21909345, 21909346, 21909347, 21909348, 21909349, 21909350, 21909351, 21909352, 21909353, 21909354, 21909355, 21909356, 21909357, 21909358, 21909359, 21909360, 21909361, 21909362, 21909363, 21909364, 21909365, 21909366, 21909367, 21909368, 21909369, 21909370, 21909371, 21909372, 21909373, 21909374, 21909375, 21909376, 21909377, 21909378, 21909379, 21909380, 21909381, 21909382, 21909383, 21909384, 21909385, 21909386, 21909387, 21909388, 21909389, 21909390, 21909391, 21909392, 21909393, 21909394, 21909395, 21909396, 21909397, 21909398, 21909399, 21909400, 21909401, 21909402, 21909403, 21909404, 21909405, 21909406, 21909407, 21909408, 21909409, 21909410, 21909411, 21909412, 21909413, 21909414, 21909415, 21909416, 21909417, 21909418, 21909419, 21909420, 21909421, 21909422, 21909423, 21909424, 21909425, 21909426, 21909427, 21909428, 21909429, 21909430, 21909431, 21909432, 21909433, 21909434, 21909435, 21909436, 21909437, 21909438, 21909439, 21909440, 21909441, 21909442, 21909443, 21909444, 21909445, 21909446, 21909447, 21909448, 21909449, 21909450, 21909451, 21909452, 21909453, 21909454, 21909455, 21909456, 21909457, 21909458, 21909459, 21909460, 21909461, 21909462, 21909463, 21909464, 21909465, 21909466, 21909467, 21909468, 21909469, 21909470, 21909471, 21909472, 21909473, 21909474, 21909475, 21909476, 21909477, 21909478, 21909479, 21909480, 21909481, 21909482, 21909483, 21909484, 21909485, 21909486, 21909487, 21909488, 21909489, 21909490, 21909491, 21909492, 21909493, 21909494, 21909495, 21909496, 21909497, 21909498, 21909499, 21909500, 21909501, 21909502, 21909503, 21909504, 21909505, 21909506, 21909507, 21909508, 21909509, 21909510, 21909511, 21909512, 21909513, 21909514, 21909515, 21909516, 21909517, 21909518, 21909519, 21909520, 21909521, 21909522, 21909523, 21909524, 21909525, 21909526, 21909527, 21909528, 21909529, 21909530, 21909531, 21909532, 21909533, 21909534, 21909535, 21909536, 21909537, 21909538, 21909539, 21909540, 21909541, 21909542, 21909543, 21909544, 21909545, 21909546, 21909547, 21909548, 21909549, 21909550, 21909551, 21909552, 21909553, 21909554, 21909555, 21909556, 21909557, 21909558, 21909559, 21909560, 21909561, 21909562, 21909563, 21909564, 21909565, 21909566, 21909567, 21909568, 21909569, 21909570, 21909571, 21909572, 21909573, 21909574, 21909575, 21909576, 21909577, 21909578, 21909579, 21909580, 21909581, 21909582, 21909583, 21909584, 21909585, 21909586, 21909587, 21909588, 21909589, 21909590, 21909591, 21909592, 21909593, 21909594, 21909595, 21909596, 21909597, 21909598, 21909599, 21909600, 21909601, 21909602, 21909603, 21909604, 21909605, 21909606, 21909607, 21909608, 21909609, 21909610, 21909611, 21909612, 21909613, 21909614, 21909615, 21909616, 21909617, 21909618, 21909619, 21909620, 21909621, 21909622, 21909623, 21909624, 21909625, 21909626, 21909627, 21909628, 21909629, 21909630, 21909631, 21909632, 21909633, 21909634, 21909635, 21909636, 21909637, 21909638, 21909639, 21909640, 21909641, 21909642, 21909643, 21909644, 21909645, 21909646, 21909647, 21909648, 21909649, 21909650, 21909651, 21909652, 21909653, 21909654, 21909655, 21909656, 21909657, 21909658, 21909659, 21909660, 21909661, 21909662, 21909663, 21909664, 21909665, 21909666, 21909667, 21909668, 21909669, 21909670, 21909671, 21909672, 21909673, 21909674, 21909675, 21909676, 21909677, 21909678, 21909679, 21909680, 21909681, 21909682, 21909683, 21909684, 21909685, 21909686, 21909687, 21909688, 21909689, 21909690, 21909691, 21909692, 21909693, 21909694, 21909695, 21909696, 21909697, 21909698, 21909699, 21909700, 21909701, 21909702, 21909703, 21909704, 21909705, 21909706, 21909707, 21909708, 21909709, 21909710, 21909711, 21909712, 21909713, 21909714, 21909715, 21909716, 21909717, 21909718, 21909719, 21909720, 21909721, 21909722, 21909723, 21909724, 21909725, 21909726, 21909727, 21909728, 21909729, 21909730, 21909731, 21909732, 21909733, 21909734, 21909735, 21909736, 21909737, 21909738, 21909739, 21909740, 21909741, 21909742, 21909743, 21909744, 21909745, 21909746, 21909747, 21909748, 21909749, 21909750, 21909751, 21909752, 21909753, 21909754, 21909755, 21909756, 21909757, 21909758, 21909759, 21909760, 21909761, 21909762, 21909763, 21909764, 21909765, 21909766, 21909767, 21909768, 21909769, 21909770, 21909771, 21909772, 21909773, 21909774, 21909775, 21909776, 21909777, 21909778, 21909779, 21909780, 21909781, 21909782, 21909783, 21909784, 21909785, 21909786, 21909787, 21909788, 21909789, 21909790, 21909791, 21909792, 21909793, 21909794, 21909795, 21909796, 21909797, 21909798, 21909799, 21909800, 21909801, 2190980 |

| | | | | | |
|------|-----------------------|---|--|--------------|---|
| 2465 | 85357483 (4929, 4930) | Novel Protein sim. GBank gij4508401[ref]NP_002871.1[prRAF1 - v-raf-1 murine leukemia viral oncogene homolog 1 | Contains protein domain (PF000069) - Eukaryotic protein kinase domain | oncogene | 18108392, 52644507, 52645156, 52646365, 22278994, 22278995, 35696286, 22278996, 22278998, 264259, 29331822, 29331824, 29331825, 60424269, 60432289, 29331827, 35696052, 29331828, 264907, 29331830, 52644045, 264909, 56182435, 264511, 265007, 265008, 265009, 264910, 33657402, 60433438, 55812038, 21906754, 33109954, 265010, 265011, 87168559, 264600, 265017, 265018, 285019, 18108351, 284369, 264288, 264685, 264767, 21906765, 21906767, 21906768, 55811957, 35695917, 265020, 265021, 60170615, 52644150, 33657023, 52645129, 33657109, 27486261, 27486264, 35695763, 264628, 263972, 18108374, 35695855, 264636, 264637, 60170394, 56526486, 87168518, 60432113, 264563, 264564, 264566, 264487 264369 |
| 2466 | 85681386 (4931, 4932) | Novel Protein sim. GBank (AF051098) seven transmembrane domain orphan receptor [Mus musculus] | | | |
| 2467 | 88059485 (4933, 4934) | Novel Protein sim. GBank F16601_1, partial CDS [Homo sapiens] | | UNCLASSIFIED | 56994075, 264908, 21906768, 33657023 |
| 2468 | 87614698 (4935, 4936) | Novel Protein sim. GBank DMR-N9 protein - mouse (fragment) | Contains protein domain (PF00400) - WD domain, G-beta repeat | kinase | 29331824, 52644045, 265008, 264910, 265019, 21906765, 21906768, 265021 264288, 264628 |
| 2469 | 86294397 (4937, 4938) | Novel Protein sim. GBank (AJ243460) proteophosphoglycan [Leishmania major] | | UNCLASSIFIED | 264592, 264692, 264555, 264556, 264557, 264558, 264559, 18108385, 264482 |
| 2470 | 80223831 (4939, 4940) | Novel Protein sim. GBank (AL098749) DKFZp434G153 [Homo sapiens] | | UNCLASSIFIED | 65274572, 35696286, 29331827, 285007, 264592, 33109954, 265018, 265019, 264288, 55811957, 265020, 264693, 55811576, 56182323 |
| 2471 | 91013681 (4941, 4942) | Novel Protein sim. GBank (AF151897) CGI-139 protein [Homo sapiens] | | UNCLASSIFIED | 264092, 60432049, 29331825, 60433356, 265010, 265011, 18108351, 264764, 264288, 264692, 65274620, 18108370, 18108372, 18108374, 264634, 18108385 |
| 2472 | 95060811 (4943, 4944) | Novel Protein sim. GBank gij4928747[gb]AAD34134.1[AF151897] CGI-139 protein [Homo sapiens] | | | 60424179, 65274572, 22278999, 60424269, 29331826, 265008, 60433356, 60433438, 265010, 18108351, 264448, 264288, 264687, 264689, 265021, 264692, 65274620, 60431528, 65274791, 264556, 56182323, 60432113 |
| 2473 | 95421509 (4945, 4946) | Novel Protein sim. GBank (AL049481) putative protein [Arabidopsis thaliana] | | | |
| 2474 | 94315616 (4947, 4948) | Novel Protein sim. GBank Unknown gene product [Homo sapiens] | | | 65274572, 56994075, 264259, 29331826, 60170831, 265017, 265018, 265019, 264683, 264369, 265020, 264693, 264563, 264564 |

| | | | | | |
|------|-----------------------|--|---|--------------|--|
| 2475 | 94321693 (4949, 4950) | Novel Protein sim. GBank gij1216486 (U48852) - HT protein [Cricetulus griseus] | Contains protein domain EGF-like domain | Ig1 | 264259, 29331822, 265008, 265007, 265010, 265011, 264448, 264288, 264369, 264685, 264686, 18108357, 264768, 18108362, 264693, 18108370, 18108374, 18108379, 35898423, 83373044, 18108383, 18108385, 264584, 264565, 264567 |
| 2476 | 94315618 (4951, 4952) | Novel Protein sim. GBank gij3252827 (AC004382) - Unknown gene product [Homo sapiens] | | UNCLASSIFIED | 264259, 60424269, 66714117, 264905, 265006, 264511, 265008, 265009, 264758, 265010, 265011, 18108351, 264681, 264369, 264288, 264689, 21906767, 265020, 18108374, 264639, 18108382, 83373044, 18108385, 87168518 |
| 2477 | 20718974 (4953, 4954) | | | UNCLASSIFIED | 263978 |
| 2478 | 17659165 (4955, 4956) | | | UNCLASSIFIED | 265017 |
| 2479 | 94314569 (4957, 4958) | Novel Protein sim. GBank gij1644232[dj]BAA110821 - (D67066) N-WASP [Bos taurus] | | Im7 | 56994075, 22278999, 21906754, 264682, 21906765 |
| 2480 | 95295605 (4959, 4960) | | | UNCLASSIFIED | 264905, 264907, 264765 |
| 2481 | 94718481 (4961, 4962) | Novel Protein sim. GBank gij5689469[dj]BAA83018.11 - (AB028989) KIAA1066 protein [Homo sapiens] | | collagen | 65274572, 56182575, 22278997, 264094, 264259, 29331822, 29331824, 66714117, 29331827, 35886052, 264508, 264905, 264908, 264907, 264908, 52644045, 264909, 56182435, 265008, 264910, 33657402, 55812038, 264758, 265010, 265011, 265017, 265018, 264760, 264762, 18108351, 264764, 264288, 264766, 264686, 264768, 21908768, 55811957, 265020, 264691, 264692, 264693, 264629, 55811576, 264630, 264634, 264635, 264636, 264637, 264556, 264558, 56182323, 83373044, 60432113, 22279002 |
| 2482 | 87393165 (4963, 4964) | Novel Protein sim. GBank gij321249[pj]S28407 - guanine nucleotide-exchange activator CDC25 homolog - mouse | | UNCLASSIFIED | 29331822, 29331824, 29331825, 29331827, 264508, 264905, 264509, 264906, 264907, 264908, 264511, 264591, 264768, 264693, 264631, 264632, 264636, 264638, 264639, 264563 |
| 2483 | 87731583 (4965, 4966) | | | UNCLASSIFIED | 264488, 22278995, 264093, 264095, 60432049, 60433356, 60433438, 264448, 264288, 263967, 18108370, 18108385, 18108388, 264482 |
| 2484 | 94187774 (4967, 4968) | Novel Protein sim. GBank gij728831[sp]P39188[ALU1_HUMAN - III] ALU SUBFAMILY J WARNING ENTRY III! | | kinase | 264563 |
| 2485 | 87786556 (4969, 4970) | Novel Protein sim. GBank gij1185397 (U25281) - SH3 domain binding protein [Rattus norvegicus] | | UNCLASSIFIED | 22278995, 22278996, 22278997, 22278999, 264259, 60432049, 29331824, 60432289, 29331827, 265007, 264910, 264593, 264600, 264603, 264604, 265018, 264448, 264288, 264685, 264686, 264769, 264689, 35695917, 265022, 264692, 264693, 56182323 |

| | | | | | |
|------|-----------------------|---|---|------------------|--|
| 2486 | 87748978 (4971, 4972) | Novel Protein sim. GBank gi 2662167 dbj BAA23715 - (AB007903) KIAA0443 [Homo sapiens] | | | 265017, 264555 |
| 2487 | 95343105 (4973, 4974) | Novel Protein sim. GBank gi 464559 sp P35287 RB14_RAT RAS-RELATED PROTEIN RAB-14 | Contains protein domain (PF00071) - Ras family | glycoprotein | 22278996, 35696286, 22278997, 22278998, 22278999, 264092, 264259, 29331822, 35696052, 264108, 264905, 264907, 29331830, 264909, 265006, 264511, 265008, 265009, 60433438, 21906754, 33109954, 87168559, 265018, 264681, 264288, 264667, 21906765, 21906768, 21906767, 21906768, 21906769, 35695917, 265021, 265022, 264534, 33657023, 264692, 33657109, 263972, 18108377, 35696423, 35695855, 60170394, 18108385, 56528486, 22278000, 22279002, 264563, 264482, 264565, 20281169, 18108391 |
| 2488 | 87652451 (4975, 4976) | | | UNCLASSIFIED | 264910, 264448, 264288, 264684, 264691, 264634 |
| 2489 | 82990585 (4977, 4978) | Novel Protein sim. GBank gi 4886439 emb CAB43355.1 - (AL050253) hypothetical protein [Homo sapiens] | Contains protein domain (PF00084) - Sushi domain (SCR repeat) | complementrecept | 264686, 264693, 55811576, 22279002 |
| 2490 | 88069609 (4979, 4980) | Novel Protein sim. GBank gi 2588624 (AC003083) - Rap2 interacting protein-like; similar to U73941 (PID:g1916018) [Homo sapiens] | | UNCLASSIFIED | 264907, 265008, 22279002 |
| 2491 | 91242116 (4981, 4982) | Novel Protein sim. GBank gi 728832 sp P39189 ALU2_HUMAN - IIII ALU SUBFAMILY SB WARNING ENTRY IIII | | tm7 | 264259, 29331826, 265008, 264762, 18108370, 18108376, 18108379 |
| 2492 | 95308202 (4983, 4984) | Novel Protein sim. GBank gi 3355303 (AF001549) - Unknown gene product [Homo sapiens] | | transcriptfactor | 264488, 22278998, 22278999, 29331828, 264591, 33109954, 265017, 55811150, 21906764, 21906768, 264692, 60431528, 87168518, 60432113, 22278000 |

| | | | | | |
|------|-----------------------|---|--|--------------|--|
| 2493 | 95422415 (4985, 4986) | Novel Protein sim. GBank gij4240307[idbj]BAA74932.1] - (AB020716) KIAA0909 protein [Homo sapiens] | Contains protein domain (PF01424) - R3H domain | struct | 18108394, 264887, 65274572, 56182575, 22278995, 56994075, 80432048, 29331822, 29331824, 29331825, 29331826, 29331827, 29146498, 264508, 264905, 264509, 264906, 264907, 29331830, 264909, 264510, 265006, 264511, 265007, 264512, 265008, 265009, 264910, 21808754, 265011, 264600, 265017, 265018, 264604, 264605, 265019, 55811150, 264762, 18108351, 264681, 264448, 264683, 264369, 264288, 18108355, 18108357, 264687, 21908765, 21906766, 21906767, 21906768, 21908769, 265020, 264691, 264692, 33657023, 33657349, 18108370, 18108374, 18108376, 55810764, 18108379, 65274791, 264630, 264632, 264634, 264635, 264636, 264555, 264637, 264557, 264558, 264639, 264559, 83373044, 18108385, 87168518, 60432113, 22279000, 22279002, 264482, 264566, 264486 |
| 2494 | 30783118 (4987, 4988) | | | UNCLASSIFIED | 264907, 264601 |
| 2495 | 94234551 (4989, 4990) | Novel Protein sim. GBank gij5420389[emb]CAB46680.1] - (AJ243460) proteophosphoglycan [Leishmania major] | | collagen | 263994, 22278997, 35696052, 264509, 264905, 264906, 264907, 264908, 264909, 265006, 265009, 264595, 264604, 264448, 264682, 264764, 264288, 264685, 264768, 264769, 264689, 265020, 264692, 65274620, 264629, 55810764, 35696423, 55811576, 264638, 264637, 18108385, 22279000, 264564, 264567, 264486 |
| 2496 | 80018765 (4991, 4992) | Novel Protein sim. GBank gij4808220[emb]CAB42832.1] - (AL022315) dJ117715.1 (PUTATIVE novel protein) [Homo sapiens] | | struct | 29147620, 264905, 265008, 265007, 18108348, 18108362, 18108370, 18108374, 264555, 264556, 18108381, 18108383, 18108388 |
| 2497 | 91723554 (4993, 4994) | | | UNCLASSIFIED | 52644507, 22278996, 22278999, 29331824, 29331828, 33657402, 21908754, 87168474, 265019, 264369, 264689, 21906765, 21908766, 21906767, 21908768, 265020, 33657023, 18108376, 18108387 |
| 2498 | 87724633 (4995, 4996) | Novel Protein sim. GBank gij1200503 (U47924) - B [Homo sapiens] | | UNCLASSIFIED | 29331827, 264512, 264910, 264288, 18108374, 35695855 |
| 2499 | 94685125 (4997, 4998) | Novel Protein sim. GBank gij3510234 (AC005581) - R31237_1, partial CDS [Homo sapiens] | Contains protein domain (PF000069) - kinase Eukaryotic protein kinase domain | kinase | 264909, 55812038, 264631, 264637, 264558 |

| | | | | | |
|------|-----------------------|--|---|--------------|--|
| 2500 | 94648324 (4989, 5000) | Novel Protein sim. GBank gij3881275 emb CAA21725 - (AL032655) predicted using GeneFinder, similar to Inositol monophosphatase family; cDNA EST yk255e11.5 comes from this gene [Caenorhabditis elegans] | Contains protein domain (PF00459) - Inositol monophosphatase family | transport | 52644507, 52645156, 22278995, 56994075, 35696286, 22278998, 264259, 52845080, 29331824, 29331825, 68714117, 60432289, 29331826, 29331827, 35696052, 29331828, 264508, 264509, 264510, 264512, 33657402, 60433438, 21906754, 52644296, 87188474, 87188559, 264603, 264681, 284448, 284683, 264288, 264369, 52644229, 264689, 21906765, 21906768, 21906767, 21906768, 21906769, 55811957, 35695917, 265020, 265021, 52644150, 33657023, 264693, 33657182, 35695763, 35696423, 35695855, 52644332, 83373044, 18108387, 87188518, 22279002 |
| 2501 | 94303896 (5001, 5002) | Novel Protein sim. GBank gij4929815 gb AAD34088.1 AF15183 - (AF151831) CGI-73 protein [Homo sapiens] | Contains protein domain (PF00651) - BTB/POZ domain | dna_mn_bind | 65274572, 56182575, 35696286, 22278998, 56994075, 22278997, 60432049, 264259, 29331822, 29331824, 29331826, 29331827, 35696052, 264905, 264906, 264907, 264908, 264909, 56182435, 264510, 264511, 265007, 264910, 264591, 60432229, 33657402, 60433356, 264595, 55812038, 264758, 264598, 87188474, 87188559, 264600, 264601, 264602, 265017, 264604, 265018, 264605, 265019, 18108351, 264448, 264369, 264288, 264766, 18108357, 21908765, 21908766, 21906767, 21906769, 29148629, 35695917, 264692, 33657023, 264629, 35696423, 55811576, 35695855, 264630, 264634, 264635, 264555, 264636, 264638, 264558, 60170394, 83373044, 18108385, 18108387, 87188518, 60432113, 22279002, 264566 |
| 2502 | 90993716 (5003, 5004) | Novel Protein sim. GBank gij3041847 (AC004542) - OXYSTEROL-BINDING PROTEIN-like; similar to P22059 (PID:g129308) [Homo sapiens] | Contains protein domain (PF01237) - Oxysterol-binding protein | UNCLASSIFIED | 65274572, 264907, 56182435, 265007, 264592, 264780, 18108351, 264448, 264369, 264288, 264684, 264886, 55811957, 265021, 264692, 33657109, 263973, 55811576, 264635, 264555, 264556, 264557, 264558, 56182323, 264559, 87188518, 264563, 264482 |
| 2503 | 87878345 (5005, 5006) | Novel Protein sim. GBank gij2196874 emb CAA72638 - (Y11896) BRX protein [Mus musculus] | | | 264905, 264907, 264512, 265008, 265011, 18108351, 264448, 264288, 29148627, 264693, 18108370, 18108374, 18108385 |

| | | | | | |
|------|-----------------------|--|--|--------------|--|
| 2504 | 87868706 (5007, 5008) | Novel Protein sim. GBank gi 550420 emb CAA48220 - (X88101) trg [Rattus norvegicus] | | | 264488, 52644507, 52645156, 52646842, 22278994, 264259, 52645080, 29331822, 29331824, 29331825, 29331826, 29331827, 35696052, 264906, 264808, 52644045, 265009, 60433356, 33657402, 60433438, 264595, 33109954, 87168474, 265017, 265019, 264448, 264288, 264766, 52644229, 21906765, 21906766, 21906767, 21906768, 52644150, 264682, 27486261, 27486262, 27486264, 27486265, 35695763, 35696423, 35695855, 52644332, 56182323, 18108387, 87168518, 60432113, 22279002, 264554 |
| 2505 | 8760559 (5009, 5010) | | | UNCLASSIFIED | 264605 |
| 2506 | 91232326 (5011, 5012) | Novel Protein sim. GBank gi 2137562 pir I49635 - mouse Dhm1 protein - mouse | | nuclease | 264488, 52644507, 52645156, 52646365, 65274572, 22278995, 56994075, 22278996, 22278997, 22278998, 22278999, 264259, 60432049, 29331822, 29331825, 29331826, 29331828, 264509, 56182435, 264112, 264593, 60433356, 55812038, 21906754, 265011, 265017, 265018, 265019, 264605, 264762, 18108351, 264448, 264288, 264768, 21906765, 21906766, 21906767, 21906768, 21906769, 35695917, 265020, 265021, 265022, 60170615, 33657023, 27486264, 18108379, 35695855, 264637, 83373044, 18108385, 87168518, 60432113, 22279000, 264563, 264482, 264565 |
| 2507 | 95316233 (5013, 5014) | Novel Protein sim. GBank gi 5174489 ref NP_006035.1 pKIAA - histone deacetylase 6 | Contains protein domain (PF00850) - Histone deacetylase family | histone | 264488, 263994, 264592, 264595, 264369, 264886, 264768, 35695917, 35696423, 264563 |
| 2508 | 85315505 (5015, 5016) | Novel Protein sim. GBank gi 4826433 emb CAB42889.1 - (AL031447) dJ126A5.2.1 (novel protein) (isolom 1) [Homo sapiens] | | UNCLASSIFIED | 22278995, 22278999, 60432049, 264259, 29331828, 265006, 265007, 60433438, 33657084, 265010, 265017, 265018, 265019, 18108351, 264448, 18108354, 264369, 18108359, 21906765, 21906769, 55811957, 265020, 265022, 27486261, 33657349, 18108377, 35695855, 60432113, 22279002, 264563, 264565 |
| 2509 | 87813741 (5017, 5018) | Novel Protein sim. GBank gi 263289 (U47856) - fibroin-4 [Araneus diadematus] | | UNCLASSIFIED | 265007, 265008, 18108357, 264556, 264567 |

| | | | | | | |
|------|-----------------------|---|---|--------------|--|--|
| 2510 | 85421379 (5019, 5020) | Novel Protein sim. GBank gij3293537[gbjAAC25762.1] - (AF071059) zinc finger RNA binding protein [Mus musculus] | | | | 65274572, 22278994, 22278996, 22278998, 22278999, 60432049, 264259, 29331822, 29331824, 29331825, 66714117, 60432289, 29331826, 33656970, 264908, 66712502, 265007, 264910, 60170831, 60432229, 60433356, 60433438, 21906754, 87168474, 265017, 265018, 264448, 264288, 21906767, 21906768, 21906769, 55811957, 35695917, 265020, 265022, 264691, 33857023, 264693, 65274620, 33657109, 33657182, 27486282, 33657349, 18108370, 35695855, 264555, 56182323, 83373044, 60432113, 22279002, 265017, 21906764, 265020, 264692 |
| 2511 | 87384281 (5021, 5022) | Novel Protein sim. GBank gij4323152[gbjAAD16228.1] - (AF098863) Ets-protein Spi-C [Mus musculus] | | | | |
| 2512 | 88084771 (5023, 5024) | Novel Protein sim. GBank gij4502075[refjNP_001135.1]pAMFR - autocrine motility factor receptor | Contains protein domain (PF000097) - Zinc finger, C3HC4 type (RING finger) | transport | | 22278999, 264259, 29331825, 29331826, 29146498, 264907, 264809, 265006, 265008, 264591, 60432229, 21906754, 264763, 264683, 264766, 18108357, 264689, 21906769, 264693, 18108370, 263972, 18108374, 264558, 22278900 |
| 2513 | 85357843 (5025, 5026) | Novel Protein sim. GBank gij3004657 (AF017777) - bobby sox [Drosophila melanogaster] | | UNCLASSIFIED | | 60424179, 52845156, 18108394, 22278994, 35896286, 56994075, 22278996, 29331822, 29331824, 60424269, 29331825, 29331827, 33656970, 60431735, 33657084, 87168559, 265017, 264448, 264369, 56181562, 21906766, 21906767, 21906768, 21906769, 265020, 265021, 33857023, 18108366, 33657109, 27486261, 27486262, 33657349, 18108374, 55810764, 35696423, 56182323, 264558, 18108385 |
| 2514 | 88094578 (5027, 5028) | Novel Protein sim. GBank gij2258437 (AF008197) - syncoilin [Rattus norvegicus] | | UNCLASSIFIED | | 264510 |
| 2515 | 87994509 (5028, 5030) | Novel Protein sim. GBank gij3757727[embjCAA18783] - (AL022727) dJ80119.7 (olfactory receptor-like protein (hs6M1-3)) [Homo sapiens] | Contains protein domain (PF00001) - 7 transmembrane receptor (rhodopsin family) | tm7 | | |
| 2516 | 87786908 (5031, 5032) | | | UNCLASSIFIED | | 264259, 29146498, 264905, 264288, 29148629, 35695917, 27486261, 264634 |
| 2517 | 87784968 (5033, 5034) | Novel Protein sim. GBank gij4220527[embjCAA23000] - (AL035356) putative protein [Arabidopsis thaliana] | | UNCLASSIFIED | | 264091, 29331824, 29331825, 29331826, 29331828, 35696052, 264508, 264509, 264905, 264906, 264907, 264908, 264909, 264511, 264910, 33657402, 264757, 33109954, 265017, 265018, 264605, 264760, 264762, 264763, 264766, 264768, 264769, 33657109, 33657182, 264628, 55811576, 35696423, 264631, 264634, 264637, 264638, 264639, 87168518, 22279002, 264564 |

| | | | | | |
|------|-----------------------|--|---|--------------|--|
| 2518 | 94147410 (5035, 5038) | Novel Protein sim. GBank gi 492959 gb AAD34056.1 AF15181 - (AF151819) CGI-61 protein [Homo sapiens] | Contains protein domain (PF00018) - SH3 domain | UNCLASSIFIED | 35696286, 264259, 29331822, 29331824, 29331825, 60432289, 29331828, 29331827, 35696052, 29331828, 264907, 264909, 264511, 265007, 60432229, 60433356, 60433438, 55812038, 265010, 265017, 264448, 264288, 264689, 21906768, 21906769, 265022, 52644150, 264693, 18108370, 263972, 264555, 56182323, 83373044, 18108385, 60432113, 264088 264259, 6671417, 29331826, 29331827, 29331828, 264907, 66712502, 265006, 265008, 264594, 265010, 265011, 265018, 264288, 21906769, 265020, 60431528, 55811576, 65274791, 264632, 264555, 264636, 22279002, 264564 |
| 2519 | 94326180 (5037, 5038) | Novel Protein sim. GBank gi 4826722 ref NP_005085.1 pFATP - fatty acid transport protein 4 | kinase | | 264259, 264908, 264910, 264682, 21906769, 265020, 264563 |
| 2520 | 87413235 (5039, 5040) | Novel Protein sim. GBank gi 4826722 ref NP_005085.1 pFATP - fatty acid transport protein 4 | transport | | 264259, 264908, 264910, 264682, 21906769, 265020, 264563 |
| 2521 | 95316244 (5041, 5042) | Novel Protein sim. GBank gi 5174489 ref NP_005035.1 pKIAA - histone deacetylase 6 | Contains protein domain (PF00850) - Histone deacetylase family | histone | 264488, 264489, 263994, 65274572, 22278995, 22278998, 264259, 29331822, 29331826, 264508, 264905, 264509, 264906, 264907, 66712502, 264511, 265006, 265007, 264591, 264592, 264593, 264594, 264595, 264596, 264681, 264448, 264763, 264882, 264764, 264684, 264369, 264288, 264685, 264686, 21906768, 55811957, 264692, 264693, 27486261, 18108370, 264628, 264629, 18108374, 55811576, 35696423, 35695855, 264632, 264558, 18108385, 65274727, 60432113, 264563, 264564, 264565, 264566, 264567 |
| 2522 | 87754052 (5043, 5044) | Novel Protein sim. GBank gi 4580011 gb AAD24201.1 U81002 - (U81002) TRAF4 associated factor 1 [Homo sapiens] | transport | | 264489, 22278997, 20281171, 21906754, 35695917, 263967, 263976, 263981, 20281169 |
| 2523 | 95340467 (5045, 5046) | Novel Protein sim. GBank gi 1809327 (U76374) - skm- | | | 263969 |
| 2524 | 95340469 (5047, 5048) | Novel Protein sim. GBank gi 1809327 (U76374) - skm- BOP2 [Mus musculus] | Contains protein domain (PF01753) - MYND finger | | 56994075, 22278996, 35696286, 22278997, 22278998, 22278999, 264259, 29331822, 29331824, 29331826, 29331828, 29331830, 56182435, 264512, 265008, 60170831, 33657402, 265010, 87168559, 265019, 264288, 21906765, 21906769, 35695917, 265020, 265021, 265022, 52644150, 264691, 33657023, 33657109, 27486281, 35696423, 65274791, 264559, 83373044, 56526486, 87168518, 264567 |

| | | | | | |
|------|-----------------------|--|---|----------------|--|
| 2525 | 94126928 (5049, 5050) | Novel Protein sim. GBank gij2073564 (U80223) - eukaryotic initiation factor eIF-2 alpha kinase; DGCN2 [Drosophila melanogaster] | | kinase | 264488, 22278997, 22278999, 60432049, 60432289, 29331828, 264905, 265008, 55812038, 21906754, 265019, 264369, 21906765, 21906766, 21906767, 21906769, 35695917, 265020, 265021, 33657109, 60431528, 83373044, 60432113, 22278000, 22279002, 264565 |
| 2526 | 95289404 (5051, 5052) | Novel Protein sim. GBank gij4589828[dijBAA78836.1] - (AB023209) KIAA0992 protein [Homo sapiens] | Contains protein domain (PF00238) - Ribosomal protein L14 | ribosomal prot | 60424178, 284768, 264687, 264769, 264689, 65274572, 21906767, 56182575, 21906768, 21908769, 55811957, 22278994, 22278995, 35698286, 35695917, 22278996, 22278997, 265020, 22278998, 265021, 22278999, 265022, 264690, 264691, 60432049, 264259, 264097, 33657023, 29331822, 29331824, 60432289, 29331826, 29331827, 29331828, 27486262, 264508, 264509, 264905, 264907, 18108370, 66712502, 60431528, 264828, 264909, 18108372, 18108374, 56182435, 18108376, 55810764, 55811576, 35696423, 35695855, 265006, 265007, 264512, 265008, 265009, 264634, 264635, 60431850, 264636, 264555, 264592, 60431735, 264638, 33657402, 56182323, 60433356, 60433438, 264595, 55812038, 264596, 264758, 83373044, 52646317, 18108385, 33657084, 18108387, 55811386, 65274727, 56526486, 87168518, 60432113, 265017, 22279000, 265018, 265019, 264564, 18108351, 264448, 264566, 264288, 264486, 264567, 264766 |
| 2527 | 88094580 (5053, 5054) | Novel Protein sim. GBank gij2258437 (AF008197) - syncollin [Rattus norvegicus] | | UNCLASSIFIED | 264510, 264512, 264630, 264591, 264592, 264259, 264594, 264595, 264603, 264605, 18108351, 264565, 264369, 18108354 |
| 2528 | 88078380 (5055, 5056) | Novel Protein sim. GBank gij2085786 (AC002086) - similar to zinc finger 5 protein from Gallus gallus, U51640 (PID:g1395185) [Homo sapiens] | Contains protein domain (PF00651) - BTB/POZ domain | dna_ma_bind | 55812038, 265017, 264689, 35695917, 35695763, 60431528, 60432113, 22278002 |
| 2529 | 86670926 (5057, 5058) | Novel Protein sim. GBank gij3786433 (AF098505) - similar to Arabidopsis thaliana male sterility protein 2 (SW:Q08891) [Caenorhabditis elegans] | | synthase | 264908, 264769, 265020, 265021, 18108383 |
| 2530 | 80259878 (5059, 5060) | | | | 264369, 264556 |
| 2531 | 87768931 (5061, 5062) | | | UNCLASSIFIED | 29331822, 29331824, 60432289, 264508, 264509, 264906, 265011, 264769, 21906768, 33657023, 87168518, 22279000 |
| 2532 | 87419778 (5063, 5064) | Novel Protein sim. GBank gij2864625[emb]CAA16972] - (AL021811) putative protein [Arabidopsis thaliana] | | | 264593 |
| 2533 | 87002255 (5065, 5066) | Novel Protein sim. GBank gij437181 (U02289) - GTPase-activating protein [Caenorhabditis elegans] | | UNCLASSIFIED | 264555 |

| | | | | | |
|------|-----------------------|---|-----------|--------------|--|
| 2534 | 87332322 (5087, 5068) | Novel Protein sim. GBank gij3452473 (AF084205) - serine/threonine protein kinase TAO1 [Rattus norvegicus] | | UNCLASSIFIED | 264259, 35696052, 264905, 265017, 21906769, 265020, 265022, 33657109, 22279000 |
| 2535 | 91225056 (5069, 5070) | Novel Protein sim. GBank gij4468311[embjCAB37992] - (AL031432) dJ465N24.2.1 (PUTATIVE novel protein) (isoform 1) [Homo sapiens] | | | 65274572, 35696286, 60432289, 29331828, 66712502, 265006, 60432229, 265017, 265018, 265019, 264288, 264369, 264689, 21905768, 265020, 265021, 264636, 60170394, 22279002 |
| 2536 | 94218540 (5071, 5072) | Novel Protein sim. GBank gij728836[spP39193]ALU8_HUMAN - IIII ALU SUBFAMILY SP WARNING ENTRY IIII | kinase | | 18108398, 56182575, 35696286, 22278997, 22278999, 60432049, 264259, 28331824, 29331826, 29331827, 29331828, 264905, 264511, 265009, 264910, 264596, 52646317, 18108351, 264681, 264683, 18108354, 264288, 264687, 264769, 264689, 21906765, 21906766, 21906767, 265021, 52645129, 33657109, 18108374, 18108380, 56182323, 18108381, 18108388, 87168518, 60432113, 22279000, 22279002, 264567, 18108391 |
| 2537 | 95422283 (5073, 5074) | Novel Protein sim. GBank gij4557026[refNP_003913.1]pHERC - guanine nucleotide exchange factor p532 | ubiquitin | | 65274572, 35696286, 29331822, 29331825, 29331827, 29331828, 35696052, 264906, 66712502, 264909, 265008, 265011, 264760, 264288, 264685, 35695917, 60170615, 264691, 33657023, 65274620, 33657109, 18108374, 35696423, 35695855, 264636, 264558, 60170394, 56182323, 83373044 |
| 2538 | 36853454 (5075, 5076) | | | UNCLASSIFIED | |
| 2539 | 94144818 (5077, 5078) | | | UNCLASSIFIED | 22278996, 22278999, 29331822, 29331825, 29331828, 29148499, 264908, 264112, 60170831, 87168559, 264604, 265019, 264685, 264768, 87168518, 22279000, 264565, 264566 |
| 2540 | 94218545 (5079, 5080) | Novel Protein sim. GBank gij1362647[pirjS53876 - sex-regulated protein janus A - fruit fly (Drosophila pseudoobscura)] | | UNCLASSIFIED | 22278997, 29331828, 265008, 265009, 264758, 265010, 18108351, 264683, 264288, 21906765, 35695917, 265020, 18108374, 264567 |
| 2541 | 95308238 (5081, 5082) | Novel Protein sim. GBank gij1711658[spP54797]T10_MOUSE - SER/THR-RICH PROTEIN T10 IN DGCR REGION | | UNCLASSIFIED | 264488, 264768, 264689, 264511, 20281171, 264634, 264635, 264691, 264639, 29331824, 264603, 264604, 264905, 264907, 264908, 264768 |

| | | | | | |
|------|-----------------------|--|--|--------------|---|
| 2542 | 85298162 (5083, 5084) | Novel Protein sim. GBank gi 5225320 gb AAD40850.1 AF08310 - (AF083107) sirtuin type 2 [Homo sapiens] | Contains protein domain (PF00220) - Neurohypophysial hormones, N- terminal Domain | UNCLASSIFIED | 264488, 18108394, 52646385, 52646842, 65274572, 22278994, 35698286, 22278996, 264259, 52645080, 29331822, 29331824, 29331827, 35698052, 33656970, 264907, 264909, 52644045, 264510, 265008, 264512, 265007, 265008, 265009, 264910, 60431735, 52646317, 52644286, 265010, 265011, 265018, 265019, 18108351, 264683, 264288, 264685, 264687, 52644229, 264769, 21908768, 21908767, 21908769, 52644150, 33657023, 33657109, 52645129, 33657182, 27486261, 27486264, 33657349, 35695763, 18108374, 35696423, 35695855, 264631, 264634, 264635, 264558, 83373044, 18108385, 18108387, 87168518, 264563, 264564 |
| 2543 | 94139088 (5085, 5086) | Novel Protein sim. GBank gi 5419857 emb CAB46374.1 - (AL086723) hypothetical protein [Homo sapiens] | Contains protein domain (PF00076) - RNA recognition motif (a.k.a. RRM, RBD, or RNP domain) | | 55274572, 56182575, 22278999, 264259, 29331826, 264907, 264510, 264511, 264592, 264595, 264764, 264369, 264288, 264684, 264766, 264689, 21908765, 21908767, 21908769, 60170615, 264682, 264693, 55811576, 85274791, 264636, 264558, 18108381, 60170394, 264639, 18108385, 60432113, 22279000 |
| 2544 | 84218549 (5087, 5088) | Novel Protein sim. GBank gi 2498110 sp Q63191 AEGP_RAT - APICAL ENDOSOMAL GLYCOPROTEIN PRECURSOR | Contains protein domain (PF00629) - MAM domain. | glycoprotein | 18108397, 52646365, 22278997, 264259, 60432049, 29331822, 29331825, 29331826, 29331827, 29331828, 264905, 264908, 265006, 265007, 265008, 87168559, 265017, 265018, 265019, 18108351, 264448, 264888, 264687, 264689, 21908765, 265020, 265021, 18108370, 18108374, 18108376, 18108381, 18108385, 18108387, 56526486, 22279000, 264482, 264563, 264567 |
| 2545 | 87742845 (5089, 5090) | Novel Protein sim. GBank gi 3327046 dbj BAA31591 - (AB014516) KIAA0616 protein [Homo sapiens] | | | 29331825, 264906, 265009, 60170831, 265017, 264369, 21908767, 60170815, 264692, 33657109 |
| 2548 | 88093861 (5091, 5092) | Novel Protein sim. GBank gi 2996032 (AF054586) - brain finger protein [Rattus norvegicus] | Contains protein domain (PF00097) - Zinc finger, C3HC4 type (RING finger) | UNCLASSIFIED | 29331824, 265007, 22279002 |

| | | | | | |
|------|-----------------------|---|---|--------------|--|
| 2547 | 94143868 (5093, 5094) | Novel Protein sim. GBank gi 49298607 gb AAD34084.1 AF15182 - (AF151827) CGI-69 protein [Homo sapiens] | Contains protein domain (PF00153) - Mitochondrial carrier proteins | transport | 264488, 18108394, 52646842, 18108397, 56182575, 22278995, 56994075, 22278996, 22278997, 22278999, 264259, 29331822, 29331824, 29331826, 60432289, 29331827, 35696052, 29331828, 264104, 264508, 264905, 264906, 264908, 66712502, 264909, 56182435, 265006, 265007, 264512, 265008, 265009, 60170831, 60432229, 60431735, 264594, 60433438, 21906754, 52646317, 265010, 265011, 264600, 264601, 265018, 265019, 264760, 18108351, 264682, 264448, 264288, 264369, 264684, 264686, 264687, 56181562, 264688, 264689, 21906765, 21906768, 21906767, 21906768, 29148627, 21906769, 55811957, 265020, 265021, 265022, 264690, 264691, 18108382, 264692, 264693, 27486261, 18108370, 18108374, 55810764, 55811576, 35696423, 35695855, 264635, 264636, 264555, 264637, 263981, 264557, 18108380, 264638, 56182323, 264558, 264559, 83373044, 18108385, 87168518, 22279002, 264564, 264566, 264488 |
| 2548 | 88179079 (5085, 5096) | | | UNCLASSIFIED | 264488, 18108394, 52646365, 22278994, 35696286, 56994075, 22278997, 22278999, 264259, 29331822, 29331824, 29331825, 29331826, 60432289, 29331827, 29331828, 56182435, 264511, 265007, 264512, 60433356, 87168559, 264684, 264369, 52644229, 265021, 33657023, 264692, 18108374, 52644332, 264557, 18108380, 18108381, 18108382, 18108384, 18108385, 60432113, 22279000, 22279002, 264563, 264567 |
| 2549 | 94186893 (5097, 5098) | Novel Protein sim. GBank gi 728837 sp P39194 ALU7_HUMAN - IIII ALU SUBFAMILY SQ WARNING ENTRY IIII | Contains protein domain (PF00412) - LIM domain containing proteins | struct | 55182575, 22278996, 22278997, 22278998, 22278999, 264259, 264508, 264908, 29331830, 265009, 265010, 265018, 264688, 21906764, 21906765, 21906766, 21906767, 21906769, 265020, 265021, 52644150, 264691, 18108368, 60431602, 18108376, 35696423, 56182323, 18108387, 264567, 56182575, 29331822, 264105, 264512, 18108351, 35695917, 264637, 264638 |
| 2550 | 87778584 (5099, 5100) | Novel Protein sim. GBank gi 2143886 pir 52523 - nucleoporin p62 homolog - rat (fragment) | | UNCLASSIFIED | |

| | | | | | |
|------|-----------------------|---|---|--------------|--|
| 2551 | 95308400 (5101, 5102) | Novel Protein sim. GBank gi 4337103 gb AAD18079 - (AF128758) NG28 [Homo sapiens] | Contains protein domain (PF00561) - alpha/beta hydrolase fold | UNCLASSIFIED | 18108396, 65274572, 22278995, 22278998, 22278999, 264259, 29331822, 29331824, 29331825, 60432289, 29331826, 29331827, 264905, 56182435, 265007, 60433438, 55812038, 21906754, 65274444, 265017, 265018, 264605, 265019, 264288, 21906766, 21906768, 21906769, 265020, 60170615, 264693, 33657109, 35698423, 264638, 56182323, 83373044, 22279000 |
| 2552 | 95332620 (5103, 5104) | | | UNCLASSIFIED | 56182575, 35696286, 29331824, 29331826, 35696052, 29331828, 264508, 264907, 56182435, 265008, 264591, 33109954, 264760, 55811957, 35695917, 33657023, 33657109, 18108374, 55811576, 35698423, 35695855, 56182323, 264558 |
| 2553 | 95308243 (5105, 5106) | Novel Protein sim. GBank gi 1711658 sp P54797 T10_MOUSE - SER/THR-RICH PROTEIN T10 IN DGCR REGION | | UNCLASSIFIED | 264686, 264488, 263976, 264768, 29331826, 35696052, 35696423, 264601, 264511, 264602, 264910, 264634, 264760, 264555, 264762, 264906, 264592, 264691, 264566, 264908, 264684, 264587, 264909, 264768 |
| 2554 | 87761520 (5107, 5108) | Novel Protein sim. GBank gi 728835 sp P39192 ALU5_HUMAN - IIII ALU SUBFAMILY SC WARNING ENTRY IIII | | cadherin | 22278997, 29331822, 264508, 21906769, 33657023, 33657109, 56182323 |
| 2555 | 87627551 (5109, 5110) | Novel Protein sim. GBank gi 488431 emb CAB43260.1 - (AL050084) hypothetical protein [Homo sapiens] | | nuclease | 29331824, 263972 |
| 2556 | 87645533 (5111, 5112) | Novel Protein sim. GBank gi 4106984 (AC003038) - R30923_1 [Homo sapiens] | Contains protein domain (PF00514) - Armadillo/beta-catenin-like repeats | UNCLASSIFIED | 22278998, 264509, 33657402, 264683, 264684, 264766, 264689, 33657023, 33657109, 35695855, 264558, 264567 |
| 2557 | 79437803 (5113, 5114) | | | | 264595 |
| 2558 | 87617591 (5115, 5116) | Novel Protein sim. GBank gi 119110 sp P03211 EBN1_EBV - EBNA-1 NUCLEAR PROTEIN | | UNCLASSIFIED | 22278997, 29331824, 66714117, 29331825, 264906, 264511, 265018, 284448 |
| 2559 | 88096382 (5117, 5118) | Novel Protein sim. GBank gi 4538998 emb CAB39619.1 - (AL049481) AIG1-like protein [Arabidopsis thaliana] | | UNCLASSIFIED | 22278997, 29331822, 29331828, 60433356, 265011, 264288, 284765, 264766, 264769, 21906765, 21906766, 60432113, 264462 |
| 2560 | 87894530 (5119, 5120) | Novel Protein sim. GBank gi 5051389 emb CAB44995.1 - (AL078630) 573K1.3 (mm17M1-4 (novel 7 transmembrane receptor (rhodopsin family) (olfactory receptor LIKE protein)) [Mus musculus] | Contains protein domain (PF00001) - 7 transmembrane receptor (rhodopsin family) | | |
| 2561 | 88176575 (5121, 5122) | Novel Protein sim. GBank gi 5326825 gb AAD42056.1 AF044953 - NADH:ubiquinone oxidoreductase PGIV subunit [Homo sapiens] | | UNCLASSIFIED | 22278995, 35696286, 22278996, 22278997, 22278998, 22278999, 264259, 29331822, 29331824, 29331826, 29331827, 29331828, 265007, 60432229, 87168559, 265017, 265018, 265019, 264689, 21906766, 21906769, 35695917, 265020, 33657023, 33657109, 18108374, 264634, 264558, 18108385, 87168518, 22279002 |

| | | | | | |
|------|-----------------------|--|--|------------------|--|
| 2562 | 87645539 (5123, 5124) | Novel Protein sim. GBank gij4106984 (AC003038) - R30923_1 [Homo sapiens] | | UNCLASSIFIED | 56994075, 22278996, 22278997, 22278999, 264259, 29331822, 60432289, 33657402, 60433356, 21906765, 55811957, 60170615, 33657023, 264693, 35695855, 87168518, 264488, 35696286, 22278999, 264259, 29331822, 29331824, 35696052, 264508, 264907, 264908, 264909, 52644045, 264510, 264511, 265009, 264910, 264591, 264593, 33657402, 265017, 265018, 265019, 18108351, 264686, 21906767, 21906768, 55811957, 35695917, 265020, 264691, 264693, 27486262, 264628, 18108374, 35696423, 35695855, 264632, 264634, 264635, 264639, 264558, 18108384, 87168518, 22279000, 22279002, 264482, 264563, 264565, 264566, 264486, 29331822, 265007, 265010, 265019, 264769, 55811576, 56182323 |
| 2563 | 88095497 (5125, 5126) | Novel Protein sim. GBank gij4886447[emb]CAB43371.1] - (AL050270) hypothetical protein [Homo sapiens] | | UNCLASSIFIED | |
| 2564 | 80502783 (5127, 5128) | Novel Protein sim. GBank gij1352944[sp]P47179]Y9P_YEAST - HYPOTHETICAL 118.4 KD PROTEIN IN BAT2-DAL5 INTERGENIC REGION PRECURSOR | | sulfotransferase | |
| 2565 | 85530908 (5129, 5130) | Novel Protein sim. GBank gij628012[pir]A53933 - myosin I | | UNCLASSIFIED | 66714117, 264909, 263978, 264632 |
| 2566 | 80224956 (5131, 5132) | Novel Protein sim. GBank gij628012[pir]A53933 - myosin I myr 4 - rat | Contains protein domain (PF00063) - Myosin head (motor domain) | - struct | 18108370, 35695855, 264556, 264558, 18108383 |
| 2567 | 86143590 (5133, 5134) | Novel Protein sim. GBank gij466009[sp]P34548]YJ4_CAEEL - HYPOTHETICAL 33.8 KD PROTEIN R10E11.4 IN CHROMOSOME III | | synthase | 265020, 60170815 |
| 2568 | 91233099 (5135, 5136) | Novel Protein sim. GBank gij466009[sp]P34548]YJ4_CAEEL - HYPOTHETICAL 33.8 KD PROTEIN R10E11.4 IN CHROMOSOME III | | | 60424179, 18108394, 56181686, 56994075, 22278999, 264490, 264259, 29331822, 56182181, 29331824, 60424269, 29331825, 29331826, 29331828, 264509, 29331830, 265007, 265008, 265009, 33657402, 265010, 265011, 265019, 264448, 264683, 264288, 18108354, 264769, 21906766, 21906767, 35695917, 265021, 33657023, 18108362, 33657109, 33657182, 35695763, 60431528, 55810764, 18108379, 83373044, 18108385, 60432113, 264482 |

| | | | | | |
|------|-----------------------|---|--|-------------------|--|
| 2569 | 95313764 (5137, 5138) | Novel Protein sim. GBank gi 2599560 gb AAB84166.1 - (AF029674) basic leucine zipper protein LZIP [Homo sapiens] | Contains protein domain (PF00170) - bZIP transcription factor | transcript factor | 18108394, 56182575, 56181886, 22278995, 22278996, 56994075, 35696286, 22278997, 22278998, 22278999, 264259, 60432049, 29331822, 56182181, 29331824, 29331825, 60432289, 29331826, 29331827, 29331828, 35696052, 264905, 264906, 56182435, 265008, 265009, 264910, 60432229, 264592, 60433356, 60433438, 21908754, 87168559, 265017, 265018, 265019, 264682, 264448, 264288, 21908765, 21906768, 21906767, 21908768, 29148627, 21906769, 35695917, 265021, 265022, 52644150, 264690, 264691, 264692, 264693, 65274620, 263967, 35695763, 20281089, 263974, 18108374, 55810764, 35696423, 35695855, 264558, 18108381, 56182323, 18108382, 83373044, 18108385, 56526486, 87168518, 22279000, 22279002, 264566 |
| 2570 | 94136754 (5139, 5140) | Novel Protein sim. GBank gi 4758954 ref NP_004567.1 pPPP2 - protein phosphatase 2 (formerly 2A), regulatory subunit B (PR 52), beta isoform | Contains protein domain (PF01240) - Protein phosphatase 2A regulatory subunit PR55 | phosphatase | 22278996, 29331822, 29331824, 66714117, 29331825, 60432289, 29331827, 35696052, 264907, 264510, 265007, 265009, 264758, 33109954, 285019, 264686, 264689, 265020, 265021, 264691, 264692, 264693, 18108374, 35695855, 264634, 56182323, 264639, 60170394, 83373044, 22279002, 264482 |
| 2571 | 87733750 (5141, 5142) | Novel Protein sim. GBank gi 732218 sp P34609 YO60_CAEEL - HYPOTHETICAL 128.6 KD PROTEIN ZK1098.10 IN CHROMOSOME III | | stud | 264508, 264905, 264907, 264628, 18108351, 264555, 264556, 264557, 264558, 264559 |
| 2572 | 87627560 (5143, 5144) | Novel Protein sim. GBank gi 4884319 emb CAB43260.1 - (AL0500084) hypothetical protein [Homo sapiens] | | UNCLASSIFIED | 35696286, 29331827, 35696052, 264100, 264104, 264110, 264592, 21906754, 29148627, 29148629, 263972, 263974, 18108374, 263976, 35695855, 60170394, 264559, 18108385 |

| | | | | | |
|------|-----------------------|--|---|--------------|--|
| 2573 | 95313929 (5145, 5146) | Novel Protein sim. GBank gi 399138 sp P02745 C1QA_HUMAN - COMPLEMENT C1Q SUBCOMPONENT, A CHAIN PRECURSOR | Contains protein domain (PF00336) - C1q domain | complement | 264488, 60424179, 65274572, 56182575, 56181686, 22278995, 58994075, 22278997, 60432049, 264259, 29331822, 29331824, 29331825, 60432289, 29331826, 29331827, 29331828, 264104, 264107, 264508, 264906, 29331830, 264909, 264510, 265006, 264512, 265008, 265009, 264910, 264591, 264592, 60432229, 264593, 60433356, 264594, 60433438, 264595, 55812038, 264759, 21906754, 87188474, 265010, 265011, 87168559, 265017, 265018, 265019, 264761, 264762, 264763, 264764, 264389, 264288, 264685, 264766, 264686, 264687, 264688, 264769, 56181562, 264689, 21906765, 21906766, 21906767, 29148527, 21906768, 21906789, 265020, 265021, 265022, 60170615, 264690, 52644150, 264691, 264692, 33657023, 65274620, 18108365, 18108368, 27486265, 60431602, 264629, 60431528, 263976, 65274791, 35695855, 20281071, 60431850, 264637, 264638, 264558, 264639, 58182323, 60170394, 83373044, 18108384, 87168518, 60432113, 264482, 264564, 264565, 264566, 264567, 22278995, 264259, 60432289, 29331827, 29331828, 33656970, 264908, 265008, 264910, 264591, 33657402, 265018, 265019, 264448, 264764, 264389, 264288, 18108357, 21906765, 21906766, 21906768, 55811857, 60170615, 264891, 33657023, 264693, 33657109, 33657182, 27486261, 27486264, 33657349, 264636, 264555, 83373044, 18108385, 264482 |
| 2574 | 94746814 (5147, 5148) | Novel Protein sim. GBank gi 3334982 (AC005306) - R27216_1 [Homo sapiens] | Contains protein domain (PF00651) - BTB/POZ domain | UNCLASSIFIED | |
| 2575 | 87754408 (5149, 5150) | Novel Protein sim. GBank gi 4929729 gb AAD34125.1 AF15188 - (AF151888) CGI-130 protein [Homo sapiens] | | UNCLASSIFIED | 264910, 264601, 264683, 264689, 264080 |
| 2576 | 95357881 (5151, 5152) | Novel Protein sim. GBank gi 4680881 gb AAD27730.1 AF13295 - (AF132955) CGI-21 protein [Homo sapiens] | Contains protein domain (PF00442) - Ubiquitin carboxyl-terminal hydrolases family 2 | ubiquitin | 264259, 35696052, 264908, 60433438, 264681, 18108351, 264288, 52644150, 264628, 35696423 |
| 2577 | 86996621 (5153, 5154) | Novel Protein sim. GBank gi 4337103 gb AAD18079 - (AF129756) NG26 [Homo sapiens] | | UNCLASSIFIED | 29331825, 265018, 265019, 264685 |
| 2578 | 87786941 (5155, 5156) | | | UNCLASSIFIED | 264488, 264906, 264908, 264910, 264596, 264603, 264604, 264605, 264766, 21906769, 264628, 264630, 264634, 264639, 264563 |
| 2579 | 87282879 (5157, 5158) | | | UNCLASSIFIED | 29331822, 29331824, 264767 |

| | | | | | |
|------|-----------------------|---|---|------------------|--|
| 2580 | 88166788 (5159, 5160) | Novel Protein sim. GBank gi 2588628 (AC003080) - Similar to KIAA0289; 60% similarity to AB002297 (PID:g2224539) [Homo sapiens] | | | 265007, 265018, 264762 |
| 2581 | 87899048 (5181, 5182) | Novel Protein sim. GBank gi 4406642 (gb AAD20049) - (AF131809) Unknown [Homo sapiens] | Contains protein domain (PF00595) - PDZ domain (Also known as DHR or GLGF). | collagen | 56994075, 29331824, 29331826, 29331828, 264905, 60433356, 60433438, 264758, 87168559, 21908769, 265022, 35695855, 263981 |
| 2582 | 87786788 (5163, 5164) | Novel Protein sim. GBank gi 2739367 (AC002505) - putative phosphatidylinositol-4-phosphate 5-kinase [Arabidopsis thaliana] | | eph | 264488, 264907, 264908, 264910, 264764, 264684, 264766, 264636, 264555, 264565 |
| 2583 | 91220850 (5165, 5166) | Novel Protein sim. GBank gi 4378112 (emb CAA16521.1) - (AL021578) dJ453C12.2 (similar to transcription factor RBP-L) [Homo sapiens] | Contains protein domain (PF00047) - Immunoglobulin domain | transcriptfactor | 56181686, 264259, 264510, 264512, 264591, 264592, 264593, 264594, 264595, 264596, 264603, 264629, 55810764, 264630, 264637, 264565 |
| 2584 | 80430941 (5167, 5168) | | | UNCLASSIFIED | 264908, 264910, 264768, 264693, 18108374, 55811576, 56182323 |
| 2585 | 80436128 (5169, 5170) | Novel Protein sim. GBank gi 2736151 (AF021935) - myotonic dystrophy kinase-related Cdc42-binding kinase [Rattus norvegicus] | | kinase | 264768 |
| 2586 | 91226136 (5171, 5172) | | | | 22278998, 264259, 29331822, 29331824, 29331827, 29331828, 264906, 265007, 265009, 264591, 60433356, 33657402, 265018, 264762, 264288, 21906766, 21906767, 21906769, 265022, 264691, 83373044, 56526486, 22279002 |
| 2587 | 80430943 (5173, 5174) | | | | 264908, 265018, 264768, 264693, 55811576, 56182323 |
| 2588 | 80074385 (5175, 5176) | | | UNCLASSIFIED | 264564 |
| 2589 | 85515607 (5177, 5178) | Novel Protein sim. GBank gi 3021598 (emb CAA71415) - (Y10389) nuclear protein [Xenopus laevis] | | UNCLASSIFIED | 35696052, 264905, 264906, 264907, 264908, 264909, 265009, 265018, 264769, 35696423, 264636 |
| 2590 | 87054526 (5179, 5180) | Novel Protein sim. GBank gi 2104689 (U92793) - alpha glucosidase II, alpha subunit [Mus musculus] | Contains protein domain (PF01055) - Glycosyl hydrolases family 31 | glucoamylase | 22278995, 29331830, 265008, 265010, 265017, 264639 |
| 2591 | 94192167 (5181, 5182) | Novel Protein sim. GBank gi 5702202 (gb AAD47199.1) (AF129166) long-chain acyl-CoA synthetase 5 [Homo sapiens] | | eph | 264259, 29331822, 264106, 264906, 56182435, 265007, 265008, 33109954, 264448, 55811957, 265020, 18108370, 55811576, 22279002 |

| | | | | | |
|------|-----------------------|---|---|----------------|---|
| 2592 | 95332648 (5183, 5184) | Novel Protein sim. GBank gi 3024998 sp Q60936 YAB1_MOUSE - HYPOTHETICAL HEART PROTEIN | | transport | 18108397, 56182575, 35696286, 56994075, 264259, 29331822, 29331824, 29331828, 60432289, 29331827, 29331828, 264906, 264909, 265007, 285008, 264910, 60432229, 264594, 60433356, 60433438, 55812038, 18108348, 21906754, 265011, 87168559, 265017, 265019, 264764, 264369, 264288, 264766, 265021, 60170615, 33657023, 33657109, 264628, 35696423, 35695855, 264557, 264638, 60170394, 56182323, 83373044, 56526486, 87168518, 264563, 264482, 264565 |
| 2593 | 87754416 (5185, 5186) | Novel Protein sim. GBank gi 4929729 gb AAD34125.1 AF15188 - (AF15188) CGI-130 protein [Homo sapiens] | | tm7 | 22278999, 29331825, 264758, 21906754, 52646317, 265010, 18108351, 264288, 264369, 21906768, 264693, 18108370, 264637, 264638, 264482 |
| 2594 | 95305758 (5187, 5188) | Novel Protein sim. GBank gi 4929587 gb AAD34054.1 AF15181 - (AF15181) CGI-59 protein [Homo sapiens] | | UNCLASSIFIED | 264488, 18108398, 56182575, 35696286, 22278997, 264093, 264259, 29331822, 29331825, 66714117, 29331826, 264905, 264909, 52644045, 56182435, 264510, 264512, 265007, 264757, 21906754, 87168474, 265017, 264760, 264448, 264764, 264288, 264766, 264689, 21906768, 33657109, 263975, 263977, 264634, 264556, 60170394, 56182323, 56526486, 264482, 264583, 264564, 264566, 264567 264692 |
| 2595 | 79561676 (5189, 5190) | Novel Protein sim. GBank gi 4309681 gb AAD15478 - (AC006930) R33423.1 [Homo sapiens] | | UNCLASSIFIED | 22278999, 264259, 265018, 264448, 265021, 60431528 |
| 2596 | 87538637 (5191, 5192) | | | UNCLASSIFIED | 264905, 264509, 264908, 264762, 264766, 35695917, 35695855, 264635, 264636, 83373044, 264488 |
| 2597 | 94784089 (5193, 5194) | | | UNCLASSIFIED | 22278998, 264259, 29331824, 87168474, 264683, 21906766, 35695917, 264691, 33657023, 33657109, 18108370, 18108374, 264564, 264565 |
| 2598 | 86094948 (5195, 5196) | Novel Protein sim. GBank gi 1001351 dbj BAA10838 - (D64006) hypothetical protein [Synechocystis sp.] | | UNCLASSIFIED | 264766, 264769, 21906768, 33657182, 35695763, 18108370, 18108374, 264635, 264636, 56526486, 22279000, 264568 |
| 2598 | 87642889 (5197, 5198) | Novel Protein sim. GBank gi 3941737 (AF109719) - BAT2 [Mus musculus] | | MHC | 35696286, 264093, 264288, 21906769, 35696423, 35695855 |
| 2600 | 87787846 (5199, 5200) | Novel Protein sim. GBank gi 4263521 gb AAD15347 - (AC004044) putative WD-repeat protein [Arabidopsis thaliana] | Contains protein domain (PF00400) - WD domain, G-beta repeat | kinasereceptor | |

| | | | | | |
|------|-----------------------|--|--|--------------|--|
| 2601 | 91243070 (5201, 5202) | Novel Protein sim. GBank gi 728837 sp P39194 ALU7_HUMAN - IIII ALU SUBFAMILY SQ WARNING ENTRY IIII | | kinase | 56182575, 22278998, 264259, 29331822, 29331824, 60432289, 29331826, 29331827, 264908, 265007, 265008, 264591, 60433356, 33857402, 60433438, 21906754, 265011, 265018, 265019, 18108351, 264448, 264369, 21906769, 265020, 60170615, 264693, 33857109, 18108370, 18108376, 56182323, 18108381, 18108385, 22279002, 264563, 60433438, 21906754, 87168559, 264601, 264369, 264288, 21906767 |
| 2602 | 88180022 (5203, 5204) | Novel Protein sim. GBank gi 4406632 gb AAD200471 - (AF131801) Unknown [Homo sapiens] | | UNCLASSIFIED | 264488, 65274572, 22278995, 22278996, 56994075, 22278997, 22278998, 22278999, 60432049, 264259, 29331822, 29331824, 60432289, 29331826, 29331827, 29331828, 35896052, 56182435, 264113, 265008, 265009, 60433356, 264757, 60433438, 264759, 33657084, 87168474, 265010, 265011, 87168559, 265017, 265018, 265019, 264448, 264683, 18108354, 264288, 264767, 264689, 21906765, 21906766, 21906767, 21906768, 21906769, 55811957, 265020, 265021, 265022, 60170615, 264691, 33657023, 264693, 33657109, 27486262, 18108374, 35898423, 85274791, 35895855, 264555, 264636, 264637, 56182323, 83373044, 56526486, 87168518, 60432113, 22279000 |
| 2604 | 94676601 (5207, 5208) | Novel Protein sim. GBank gi 5454030 ref NP_006468.1 pRRP2 - RAS-related on chromosome 22 | | oncogene | 264259, 35896052, 264508, 264906, 264907, 264908, 264909, 264510, 264512, 265008, 264910, 33657402, 264604, 264605, 264762, 264763, 264682, 264764, 264683, 264768, 264769, 264689, 33657023, 264693, 18108365, 264628, 35898423, 264631, 264632, 264634, 264635, 264637, 18108381, 264639, 83373044, 264555 |
| 2605 | 94316756 (5209, 5210) | Novel Protein sim. GBank gi 3628745 dbj BAA33366 - (AB013721) mitsugumin 23 [Oryctolagus cuniculus] | | UNCLASSIFIED | 22278998, 264490, 60432049, 264259, 60432289, 264909, 265008, 60433356, 60433438, 264758, 21906754, 265010, 265011, 265018, 264681, 18108351, 264288, 264766, 264685, 21906765, 21906766, 21906768, 21906769, 264691, 264692, 264693, 65274791, 264634, 264555, 264636, 22278996, 264510, 264512, 265009, 264768, 22279002, 264556 |
| 2606 | 87746406 (5211, 5212) | | | | |

| | | | | | |
|------|-----------------------|---|--|-------------------|--|
| 2607 | 87627742 (5213, 5214) | Novel Protein sim. GBank gij4826626[gb]AAD30202.1] - (AF135022) mediator [Homo sapiens] | | | 29331822, 29331825, 29331826, 29331827, 29331828, 264908, 264907, 264908, 66712502, 264828, 56182435, 55812038, 265010, 265017, 265018, 265019, 264768, 264689, 21906765, 55811957, 265020, 265022, 264692, 33657023, 264693, 33657109, 18108370, 264639, 56182323, 264509, 264907, 264908, 264592, 264758, 264631 |
| 2608 | 81734786 (5215, 5216) | Novel Protein sim. GBank gij226005 (U49973) - ORF2: function unknown [Homo sapiens] | | | |
| 2609 | 94843791 (5217, 5218) | Novel Protein sim. GBank gij3024889[sp]P56524[Y288_HUMAN - HYPOTHETICAL PROTEIN KIAA0288 (HA6116)] | Contains protein domain (PF00850) - Histone deacetylase family | histone | 264488, 65274572, 35696286, 22278997, 22278999, 60432049, 264259, 56182181, 29331824, 29331825, 29331826, 60432289, 29331827, 29331828, 264905, 264907, 60433356, 60433438, 55812038, 265011, 87168559, 265017, 265018, 264448, 264765, 264288, 264766, 264689, 21906765, 21906767, 21906769, 265020, 265021, 264691, 264692, 33657109, 27486261, 18108370, 65274791, 264636, 264556, 56182323, 18108385, 56526486 |
| 2610 | 88177664 (5219, 5220) | Novel Protein sim. GBank gij4336855[gb]AAD17989] - (AF108473) leucine-rich-domain inter-acting protein 1; LeR inter-acting protein 1; LEAP1 [Mus musculus] | | transcript factor | 18108394, 22278994, 56994075, 60432049, 264259, 29331822, 29331825, 60432289, 29331827, 264107, 264109, 264905, 56182435, 264112, 265006, 265007, 265008, 265009, 60433356, 60433438, 265011, 87168559, 265017, 264448, 264682, 264764, 264288, 265021, 33657023, 263967, 33657182, 27486261, 18108374, 263976, 55811576, 264638, 87168518, 60432113, 22278999, 265017, 264684, 21906768, 22279000 |
| 2611 | 87428890 (5221, 5222) | Novel Protein sim. GBank gij3876761[emb]CAA92994] - (Z68760) predicted using GeneFinder. Similarity to Mouse FK506-binding protein (SW:FKB3_MOUSE) [Caenorhabditis elegans] | Contains protein domain (PF00254) - FKBP-type peptidyl-prolyl cis-trans isomerases | isomerase | |
| 2612 | 87771198 (5223, 5224) | Novel Protein sim. GBank gij5679136[gb]AAD46874.1[AF16093] - (AF160934) BcDNA.LD14189 [Drosophila melanogaster] | | transport | 265009, 264910, 264759, 265017, 21906767, 18108365, 18108388, 60432113 |
| 2613 | 79481496 (5225, 5226) | | | UNCLASSIFIED | 264685 |
| 2614 | 87643948 (5227, 5228) | Novel Protein sim. GBank gij5533081[gb]AAD45009.1[AF16118] - (AF161181) P55T protein [Mus musculus] | Contains protein domain (PF00625) - Guanylate kinase | | 22278998, 29331825, 264508, 264906, 21906754, 264602, 264766, 264769, 52644229, 21906765, 33657109, 27486264, 18108370, 263972, 264555, 60432113 |
| 2615 | 87381886 (5229, 5230) | | | UNCLASSIFIED | 264768, 18108394, 264692, 264693, 264508, 264509, 264907, 264628, 264908, 264909, 18108377, 264511, 264512, 264910, 264635, 264595, 265010, 264404, 264563, 264764, 264685, 264766 |

| | | | | | |
|------|-----------------------|---|--|-----------------|--|
| 2616 | 87428895 (5231, 5232) | Novel Protein sim. GBank gl 3876761 emb CAA02994 - (Z68760) predicted using Genefinder; Similarity to Mouse FK506-binding protein (SW:FKB3_MOUSE) [Caenorhabditis elegans] | Contains protein domain (PF00254) - FKBP-type peptidyl-prolyl cis-trans isomerases | isomerase | 22278995, 22278997, 22278998, 60432049, 60432289, 264828, 60433356, 264594, 60433438, 33109954, 87188474, 265011, 265017, 265019, 264288, 264766, 21906765, 21908767, 265020, 265021, 18108376, 18108377, 18108387, 87168518, 264482, 264587 |
| 2617 | 86976888 (5233, 5234) | Novel Protein sim. GBank gi 728831 sp P39188 ALU1_HUMAN - IIII ALU SUBFAMILY J WARNING ENTRY IIII | | kinase | 265010, 265019, 264369, 264693, 55811576, 22279002 |
| 2618 | 91231862 (5235, 5236) | Novel Protein sim. GBank gl 3319282 (AF049103) - Huntingtin interacting protein [Homo sapiens] | Contains protein domain (PF00397) - WW domain | UNCLASSIFIED | 264489, 22278996, 264490, 264259, 29331822, 264102, 264509, 264908, 264907, 66712502, 29331830, 265008, 264910, 265009, 60433356, 60433438, 264758, 21906754, 265011, 87168559, 265017, 265018, 264369, 264288, 264766, 264768, 264689, 21906765, 21906766, 21906767, 35695917, 265020, 265022, 33657023, 264692, 33657109, 264628, 18108374, 35695855, 18108381, 83373044, 18108385, 18108388, 56526486, 264563 |
| 2619 | 87694000 (5237, 5238) | Novel Protein sim. GBank gl 2431772 (U66411) - putative type III alcohol dehydrogenase [Drosophila melanogaster] | Contains protein domain (PF00465) - Iron-containing alcohol dehydrogenases | - dehydrogenase | 264259, 60432289, 60433438, 21906754, 264369, 60432113, 264566 |
| 2620 | 95314841 (5239, 5240) | Novel Protein sim. GBank gl 4322567 gb AAD16097 - (AF090436) dachshund variant 1 [Mus musculus] | Contains protein domain (PF00628) - PHD-finger | UNCLASSIFIED | 52644507, 52645156, 52646842, 65274572, 22278995, 56994075, 35696286, 22278999, 60432049, 264259, 52645080, 29331822, 29331824, 29331825, 29331826, 29331828, 35696052, 264907, 66712502, 265008, 60433356, 33657402, 52646317, 21906754, 87168474, 265010, 265017, 265018, 264448, 264369, 264288, 264687, 264768, 52644229, 264688, 264689, 21906765, 21906768, 35695917, 52644150, 264692, 33657109, 35695763, 35696423, 264556, 52644332, 18108382, 83373044, 18108385, 18108387, 65274727, 87168518, 60432113, 22279002 |
| 2621 | 80253495 (5241, 5242) | | | | 264594, 264636 |
| 2622 | 81780380 (5243, 5244) | Novel Protein sim. GBank gi 4557341 ref NP_001174.1 pATP6 - ATPase, H ⁺ transporting, lysosomal subunit 1; vacuolar proton pump, H-ATPase subunit | | | 264488, 264908, 264907, 264908, 264512, 265007, 264758, 35695917, 264634, 264636, 264563, 264482 |
| 2623 | 91633306 (5245, 5246) | Novel Protein sim. GBank gl 3880355 emb CAB05209 - (Z62285) predicted using Genefinder [Caenorhabditis elegans] | | UNCLASSIFIED | 29331824, 35696052, 265007, 265010, 264288, 29148629 |

| | | | | | |
|------|-----------------------|--|---|--------------|--|
| 2624 | 91639308 (5247, 5248) | Novel Protein sim. GBank gij3880355[emb]CAB05299] - (Z82285) predicted using Genefinder [Caenorhabditis elegans] | | UNCLASSIFIED | 56181686, 22278996, 22278997, 22278998, 22278999, 264259, 29331822, 29331824, 56182181, 29331825, 60432289, 29331828, 35698052, 29146489, 66712502, 52644045, 265007, 265008, 60433356, 33109954, 21906754, 265010, 265011, 265019, 264448, 264288, 21906765, 21906766, 21906767, 29148629, 35695917, 265021, 265022, 27486285, 18108370, 60431528, 55811576, 35695855, 56182323, 18108385, 87168518, 22279002, 18108391 |
| 2625 | 86452068 (5249, 5250) | Novel Protein sim. GBank gij2887429[dbj]BAA24857] - (AB007887) KIAA0427 [Homo sapiens] | | UNCLASSIFIED | 264091, 264511, 263981 |
| 2626 | 16533787 (5251, 5252) | Novel Protein sim. GBank gij487416 (L20302) - actin filament protein [Gallus gallus] | | struct | 265008 |
| 2627 | 87636823 (5253, 5254) | Novel Protein sim. GBank gij88462[pir]J27307 - proline-rich phosphoprotein (gene PRH1, Db allele) - human | | UNCLASSIFIED | 22278996, 265007, 265009, 264448, 21906767, 265021, 264558, 87168518 |
| 2628 | 94848254 (5255, 5256) | Novel Protein sim. GBank gij3123552[emb]CAA18609] - (AL022578) dJ393P12.2 (hypothetical Proline-rich protein KIAA0269 LIKE) [Homo sapiens] | | UNCLASSIFIED | 22278997, 22278999, 264259, 60432049, 29331822, 29331824, 29331825, 29331827, 35696052, 29331828, 284907, 264909, 265008, 264591, 60433356, 60433438, 265017, 265018, 264389, 264288, 18108357, 21906765, 21908768, 265022, 63274791, 264638, 18108387, 87168518, 22279002 |
| 2629 | 87376490 (5257, 5258) | Novel Protein sim. GBank gij4929595[gb]AAD34058.1[AF15182] - (AF151821) CGI-63 protein [Homo sapiens] | | synthase | 29331825, 29331826, 264102, 265006, 264766, 35695917, 264891, 33657023, 263972, 18108374, 22279000 |
| 2630 | 78188364 (5259, 5260) | Novel Protein sim. GBank gij321605[pir]JQ11161 - Gag protein - Visna virus (strain EV1) | Contains protein domain (PF00009) - Zinc finger, CCHC class | UNCLASSIFIED | 264636, 18108385 |
| 2631 | 94845909 (5261, 5262) | | | | 52644507, 52645156, 52646365, 52646842, 22278994, 22278995, 35696286, 56994075, 22278997, 22278999, 264259, 52645080, 29331822, 29331824, 29331825, 29331826, 29331827, 29331828, 35696052, 33656970, 264905, 264509, 264907, 264908, 264511, 264512, 265007, 265008, 284910, 52646317, 33657084, 52644296, 265010, 87168559, 265017, 265018, 265019, 264760, 264762, 264448, 264288, 264368, 264766, 264768, 52644229, 21906764, 21906765, 21906766, 21906767, 21906769, 35695917, 265020, 52644150, 33657023, 52645129, 33657109, 33657182, 27486261, 27486262, 27486265, 33657349, 35695763, 35696423, 65274791, 35695855, 264634, 264637, 52644332, 56182323, 60432113, 264566, 264486 |
| 2632 | 38730414 (5263, 5264) | | | | 264685 |

| | | | | |
|------|-----------------------|--|--|--|
| 2633 | 95011617 (5265, 5266) | Novel Protein sim. GBank gij113954[idb]BAA1088g] - (D84009) seizure-related gene product 6 type 2 precursor [Mus musculus] | Contains protein domain (PF00084) - Sushi domain (SCR repeat) | 22278995, 22278997, 22278998, 264259, 29331822, 29331824, 29331825, 29331827, 264508, 264906, 265006, 265007, 265008, 265009, 55812038, 33857084, 55811386, 265010, 265011, 87188559, 265018, 265019, 264683, 264288, 264886, 29148829, 33657023, 264693, 33657182, 35695763, 55811576, 264639, 56182323, 83373044, 18108385, 56526486, 87168518, 22279000, 22279002, 264565 |
| 2634 | 87330921 (5267, 5268) | Novel Protein sim. GBank gij5441611[emb]CAB46854.1] - (AJ388555) hypothetical protein [Canis familiaris] | UNCLASSIFIED | 29331826, 263972, 264089 |
| 2635 | 86623144 (5269, 5270) | Novel Protein sim. GBank gij4680663[gb]AAD27721.1[AF132946] CGI-12 protein [Homo sapiens] | | 22278997, 264259, 29331824, 66714117, 29331827, 29331828, 264907, 33657084, 265017, 265018, 264448, 264288, 21906766, 21906767, 21906768, 29148829, 18108376, 55811576, 35695855, 87168518, 22279000, 264568, 29331822, 29331828, 265006, 60170831, 264681, 264765, 264685, 29148827, 21906769, 29148784, 265022, 60170815, 264635, 18108385, 56526486, 22279002, 264567 |
| 2636 | 87260534 (5271, 5272) | Novel Protein sim. GBank gij3879146[emb]CAB07846] - (Z93386) Similarity to Yeast hypothetical 52.9 kD protein (SW:P43616); cDNA EST EMBL:M89432 comes from this gene; cDNA EST EMBL:D71008 comes from this gene; cDNA EST EMBL:D73578 comes from this gene; cDNA EST EMBL:D69025 comes ... | Contains protein domain (PF01546) - Peptidase family M20/M25/M40 | |
| 2637 | 95011298 (5273, 5274) | Novel Protein sim. GBank gij4758208[ref]NP_004081.1[PDUSP - dual specificity phosphatase 3 (vaccinia virus phosphatase VH1-related)] | Contains protein domain (PF00782) - Dual specificity phosphatase, catalytic domain | 264488, 264489, 52644507, 264887, 52646365, 52646842, 22278994, 22278996, 22278999, 20281171, 264259, 29331822, 52645080, 66714117, 29331825, 29331826, 29331827, 35696052, 29331828, 29146498, 264508, 264905, 264906, 264907, 264908, 264909, 264511, 264512, 264910, 264591, 264592, 60432229, 264593, 264594, 33657402, 60433356, 264757, 60433438, 264596, 264758, 52646317, 21906754, 52644296, 265010, 264600, 264602, 264603, 264605, 264761, 264762, 264681, 264448, 264764, 264765, 264288, 264766, 264686, 264768, 264687, 264769, 21906765, 21906766, 21906767, 21906768, 21906769, 35695917, 265021, 60170615, 264691, 33657023, 264692, 264693, 65274820, 27486264, 18108370, 264628, 264629, 18108374, 35696423, 35695855, 264632, 264634, 264635, 264637, 264638, 52644332, 264639, 264558, 83373044, 60432113, 264564, 264565, 264566, 264486, 264567 |

| | | | | | |
|------|-----------------------|--|--|--------------|--|
| 2638 | 94326733 (5275, 5276) | Novel Protein sim. GBank gi 49298899 gb AAD34105.1 AF15186 - (AF15186) CGI-110 protein [Homo sapiens] | Contains protein domain (PF00076) - RNA recognition motif (a.k.a. RRM, RBD, or RNP domain) | UNCLASSIFIED | 60424179, 52644507, 52648842, 18108398, 55182575, 22278995, 22278995, 35696286, 22278997, 22278999, 264259, 60432049, 29331822, 60424269, 29331826, 35696052, 29146498, 284905, 52644045, 56182435, 60433358, 33657402, 55812038, 55811386, 265019, 264288, 264769, 52644229, 56181562, 29148627, 29148629, 55811957, 29148784, 35695917, 265021, 52644150, 33657023, 65274620, 33657109, 35695763, 18108374, 55810784, 35696423, 55811576, 35695855, 60431850, 56182323, 60432113, 264404 |
| 2639 | 85361346 (5277, 5278) | Novel Protein sim. GBank gi 2190007 gb BAA20355 - (AB004109) phosphatidylserine synthase II [Cricketulus griseus] | | synthase | 264488, 29331825, 35696052, 264508, 264509, 264909, 264512, 33657402, 60433438, 264758, 85658542, 264600, 265020, 265021, 33657109, 264628, 35696423, 264555, 264639, 264563, 264564, 264565, 264566, 264486 |
| 2640 | 87781330 (5279, 5280) | Novel Protein sim. GBank gi 3158516 (AF087617) - contains similarity to chromo (chromatin organization modifier) domains (Pfam: chrom.hmm, score: 17.78 and 27.94) and to helicases conserved C-terminal domain (Pfam: helicase_C.hmm, score: 87.00) [Caenorhabditis elegans] | Contains protein domain (PF00271) - Helicases conserved C-terminal domain | helicase | 29331822, 29331826, 264906, 33108954, 265017, 265019, 21906788, 35695763, 264636, 264637, 18108387 |
| 2641 | 11669834 (5281, 5282) | Novel Protein sim. GBank gi 2564955 (AF030001) - unknown [Mus musculus] | | UNCLASSIFIED | 264828 |
| 2642 | 87412575 (5283, 5284) | Novel Protein sim. GBank gi 4490304 emb CAB38785.1 - (AL036676) putative protein [Arabidopsis thaliana] | | UNCLASSIFIED | 264259, 29331822, 33657402, 265019, 264369, 264691, 264634, 56526486, 22279002 |
| 2643 | 87643981 (5285, 5286) | Novel Protein sim. GBank gi 3789797 gb AAC67502.1 - (AF059569) actin binding protein MAYVEN [Homo sapiens] | Contains protein domain (PF00270) - DEAD/DEAH box helicase | helicase | 22278997, 264259, 29148499, 56182435, 264910, 265010, 18108351, 264682, 264683, 264369, 264684, 264685, 264686, 29148627, 264690, 33657109, 18108370, 263973, 18108374, 264634, 264557, 264558, 18108385, 264482 |
| 2644 | 88177671 (5287, 5288) | Novel Protein sim. GBank gi 3789797 gb AAC67502.1 - (AF059569) actin binding protein MAYVEN [Homo sapiens] | Contains protein domain (PF00651) - BTB/POZ domain | nucl_rept | 264107, 264687 |
| 2645 | 17277228 (5289, 5290) | Novel Protein sim. GBank gi 1708722 sp P49749 EVX2_MOUSE - HOMEBOX EVEN SKIPPED HOMOLOG PROTEIN 2 (EVX-2) | | UNCLASSIFIED | 265007 |
| 2646 | 94148542 (5291, 5292) | | | UNCLASSIFIED | 264908, 264687, 264632, 83373044 |

| | | | | | |
|------|-----------------------|--|--|-------------------|--|
| 2647 | 91212878 (5283, 5284) | | | UNCLASSIFIED | 56182575, 22278996, 35696286, 22278998, 264259, 29331822, 56182181, 29331825, 60424269, 60432289, 35696052, 66712502, 264908, 265007, 55812038, 33109954, 21906754, 33657084, 265019, 264448, 264288, 56181562, 21906765, 21906766, 21906768, 21906769, 35695917, 265020, 265021, 52644150, 264693, 33657109, 33657349, 60431528, 18108374, 55810764, 35696423, 56182323, 60432113, 22279002, 264584 |
| 2648 | 87600587 (5295, 5296) | | | | 29146498, 56182435, 33109954, 265011, 264682, 55811957, 35695917, 264690, 263976, 18108377, 35696423, 60432113 |
| 2649 | 94128783 (5287, 5288) | Novel Protein sim. GBank gi 3041852 (AC004539) - unknown function; similar to Y09105 (PID:g1666171) [Homo sapiens] | | UNCLASSIFIED | 56182575, 35696286, 22278998, 29331824, 29331826, 60432289, 66712502, 56182435, 60170831, 60432229, 33657402, 33109954, 21906754, 265017, 264686, 264688, 21906765, 21906768, 60170815, 264693, 263987, 18108370, 263976, 60170394, 60432113, 22279002, 264563 |
| 2650 | 87287533 (5289, 5300) | Novel Protein sim. GBank gi 5360271 dbj BAA81908.1 - (AB029335) HrPET-3 [Halocynthia roretzi] | | | 264685 |
| 2651 | 88088745 (5301, 5302) | Novel Protein sim. GBank gi 4240225 dbj BAA74891.1 - (AB020675) KIAA0888 protein [Homo sapiens] | Contains protein domain (PF00054) Laminin G domain | | 29331824, 29331826, 29331827, 265007, 55812038, 21906754, 18108368, 18108384, 22279002, 264567 |
| 2652 | 10343125 (5303, 5304) | Novel Protein sim. GBank gi 4493959 emb CAB11123.2 - (Z98551) predicted using hexExon; MAL3P6.28 (PFC0845c), Hypothetical protein, len: 167 aa; Similarity to model organism hypothetical proteins (C.elegans, D.melanogaster, S.cerevisiae & S.pombe). C.elegans protein ZK287.5 (TR.... | | UNCLASSIFIED | 264692 |
| 2653 | 87798735 (5305, 5306) | Novel Protein sim. GBank gi 4493959 emb CAB11123.2 - (Z98551) predicted using hexExon; MAL3P6.28 (PFC0845c), Hypothetical protein, len: 167 aa; Similarity to model organism hypothetical proteins (C.elegans, D.melanogaster, S.cerevisiae & S.pombe). C.elegans protein ZK287.5 (TR.... | | UNCLASSIFIED | 265018, 18108370, 18108387, 264566 |
| 2654 | 85103240 (5307, 5308) | | | | 60424179, 65274572, 56182575, 264259, 56182181, 264908, 56182435, 55811957, 35695917, 265021, 263976, 55810764, 65274791, 56182323, 83373044, 65274727 |
| 2655 | 91228018 (5309, 5310) | Novel Protein sim. GBank gi 3875272 emb CAB02861 - (Z81051) predicted using GeneFinder; similar to Zinc finger, C3HC4 type (RING finger); cDNA EST yk443h5.3 comes from this gene; cDNA EST yk443h5.5 comes from this gene [Caenorhabditis elegans] | Contains protein domain (PF000097) - Zinc finger, C3HC4 type (RING finger) | transcript factor | 56182575, 56181888, 264092, 264259, 56182181, 60432289, 264907, 33657402, 55812038, 21906754, 87168559, 265017, 264448, 264369, 264288, 21906765, 21906766, 21906767, 21906768, 33657109, 18108370, 264628, 55811576, 264556, 264639, 83373044, 56526486, 264404, 60432113 |
| 2656 | 84562601 (5311, 5312) | Novel Protein sim. GBank gi 3043718 dbj BAA25523 - (AB011169) KIAA0597 protein [Homo sapiens] | | | 264693 |

| | | | | | |
|------|-----------------------|--|---|---------------|--|
| 2657 | 52561728 (5313, 5314) | Novel Protein sim. GBank gij5689509[dbj BAA83038.1] - (AB029009) KIAA1086 protein [Homo sapiens] | | dna_rna_bind | 264693 |
| 2658 | 88062454 (5315, 5316) | Novel Protein sim. GBank gij3688089 (AC005757) - R32611_1 [Homo sapiens] | Contains protein domain (PF00560) - Leucine Rich Repeat | nucleaseinhib | 35696286, 264259, 29331822, 29331824, 29331826, 29331828, 265019, 264683, 21906768, 35695917, 264693, 35695855, 264637, 87168518, 264486, 264567 |
| 2659 | 87600755 (5317, 5318) | Novel Protein sim. GBank gij5420387[emb CAB46679.1] - (AJ243459) proteophosphoglycan [Leishmania major] | Contains protein domain (PF01426) - BAH domain | UNCLASSIFIED | 264909, 264910, 265018, 264369, 264769, 21906769, 264693, 263972, 18108388 |
| 2660 | 91718472 (5319, 5320) | Novel Protein sim. GBank gij728837[sp P39194 ALU7_HUMAN - IIII ALU SUBFAMILY SQ WARNING ENTRY IIII | Contains protein domain (PF00036) - EF hand | kinase | 264488, 65274572, 35696286, 22278998, 22278999, 264259, 29331822, 29331824, 60432289, 29331826, 35696052, 264908, 56182435, 265008, 265009, 60433356, 264594, 265010, 265018, 55811150, 18108351, 264682, 264684, 264369, 264288, 264687, 21906765, 29148784, 35695917, 60170615, 52644150, 33657023, 33657109, 35696423, 35695855, 264556, 60170394, 18108385, 22279000, 22279002 |
| 2661 | 95342817 (5321, 5322) | Novel Protein sim. GBank gij4758048[ref NP_004739.1 pCPR8 - cell cycle progression 8 protein | | glycoprotein | 60432049, 264259, 29331824, 29331825, 29331826, 29331827, 29331828, 264906, 264909, 264593, 33109954, 265010, 265017, 265018, 265019, 264760, 264448, 264369, 264288, 21906765, 21906768, 265022, 264691, 33657023, 27488252, 60431528, 18108374, 35695855, 18108388, 264482, 264555, 264556, 264558, 264486 |
| 2662 | 80228739 (5323, 5324) | Novel Protein sim. GBank gij3874714[emb CAA91263] - (Z66494) similar to choline dehydrogenase; cDNA EST yk346d5.5 comes from this gene; cDNA EST yk346d5.3 comes from this gene [Caenorhabditis elegans] | | dehydrogenase | 264906, 264909, 264757, 264758, 264767, 264691, 33657023, 264638 |
| 2663 | 87780623 (5325, 5326) | Novel Protein sim. GBank gij1389670 (U58977) - Notch homolog Scalloped wings [Lucilia cuprina] | Contains protein domain (PF00008) - EGF-like domain | oncogene | 35696286, 264509, 264595, 264288, 264685, 264686 |
| 2664 | 85518329 (5327, 5328) | Novel Protein sim. GBank gij1389670 (U58977) - Notch homolog Scalloped wings [Lucilia cuprina] | | UNCLASSIFIED | 35696286, 22278999, 29331822, 35696052, 264906, 264907, 264909, 264510, 264511, 264512, 264593, 60433438, 265019, 264681, 21906765, 21906766, 21906767, 21906768, 265020, 265022, 35696423, 35695855, 22279002, 264482, 264488 |
| 2665 | 87770662 (5329, 5330) | Novel Protein sim. GBank gij4884406[emb CAB43311.1] - (AL050190) hypothetical protein [Homo sapiens] | | UNCLASSIFIED | 29331825, 265007, 264910, 60432229, 265019, 264288, 21906767, 264558, 22279002 |
| 2666 | 87828472 (5331, 5332) | Novel Protein sim. GBank gij5108956[gb AAD39906.1 AF113615) FH1/FH2 domain-containing protein FHOS [Homo sapiens] | | nuclease | 264907, 29331830, 264681, 264683, 264288, 35695855, 264632, 264556, 264557, 264558, 284559, 264563, 264565, 264567 |
| 2667 | 87422720 (5333, 5334) | Novel Protein sim. GBank gij2500570[sp Q17533 RNPH_CAEEL - RIBONUCLEASE PH-LIKE PROTEIN B0564.1 | Contains protein domain (PF01138) - 3' exoribonuclease family | | |

| | | | | |
|------|-----------------------|--|--|---|
| 2668 | 91216716 (5335, 5336) | Novel Protein sim. GBank gij5454188[re]NP_006327.1[pZYG]- ZYG homolog | UNCLASSIFIED | 56181686, 35696288, 22278998, 22278999, 56182181, 29331824, 60424269, 29331825, 35696052, 29331828, 68712502, 56182435, 60433356, 284758, 21908754, 55811386, 265011, 87168559, 265017, 265019, 55811150, 284448, 264369, 264288, 21906765, 21906768, 21906768, 55811957, 35695917, 265020, 265021, 33657023, 264692, 33657109, 35695763, 60431528, 18108374, 35696423, 55811576, 264634, 60431850, 83373044, 18108385, 87168518, 22279000, 264563, 264564 |
| 2669 | 95415721 (5337, 5338) | Novel Protein sim. GBank gjl2147012[pi]JC4899 - proline rich protein - rat | | 264489, 264889, 21906767, 65274572, 56182575, 21906768, 29148827, 21906769, 29148629, 35696286, 35695917, 22278998, 22278998, 265021, 265022, 60170615, 52644150, 60432049, 264259, 264691, 33657023, 264692, 29331822, 29331824, 29331825, 60432289, 33657109, 29331828, 29331827, 35696052, 29331828, 29148498, 29146499, 264905, 264908, 52844045, 284909, 56182435, 35696423, 65274791, 35695855, 265006, 264910, 264635, 60432229, 264592, 264638, 56182323, 60433356, 60170394, 60433438, 264559, 264595, 55812038, 33109954, 87168559, 60432113, 265019, 264448, 264369, 264684, 264288 |
| 2670 | 87613234 (5339, 5340) | Novel Protein sim. GBank gij1723523[sp]Q10362[YDBB_SCHPO - HYPOTHETICAL 94.9 KD PROTEIN C22E12.11C IN CHROMOSOME I | Contains protein domain (PF006528) - PHD-finger | 18108370, 263974 |
| 2671 | 91214836 (5341, 5342) | Novel Protein sim. GBank gij4768277[gb]AAD29444.1[AF064255] very long-chain acyl-CoA synthetase homolog 2: VLCS-H2 [Homo sapiens] | transport | 52646842, 56994075, 284259, 29331822, 29331824, 29331825, 29331827, 33656970, 264509, 265006, 33109954, 21906754, 264682, 264288, 265021, 33657023, 33657109, 33657182, 27486261, 27486262, 27486265, 18108376, 18108385 |
| 2672 | 87399123 (5343, 5344) | Novel Protein sim. GBank gij4966346[gb]AAD34877.1[AC008341] Contains two PF01344 Kelch motif domains. [Arabidopsis thaliana] | UNCLASSIFIED Kelch motif | 264767 |

| | | | | | |
|------|-----------------------|---|---|------------------|--|
| 2673 | 87430749 (5345, 5348) | Novel Protein sim. GBank gi 5457337 emb CAB41505.2 - (AJ236876) poly(ADP-ribose) polymerase-2 [Homo sapiens] | Contains protein domain (PF00844) - Poly(ADP-ribose) polymerase catalytic region. | polymerase | 22278995, 22278996, 22278997, 22278999, 29331822, 29331824, 29331828, 35696052, 60433438, 87168474, 87168559, 265017, 265018, 265019, 264448, 21906768, 21906769, 265020, 265021, 33657109, 27486262, 35695763, 60431850, 60170394, 87168518, 264563 |
| 2674 | 94847721 (5347, 5348) | Novel Protein sim. GBank gi 4758824 ref NP_004280.1 pNRF3 - nuclear factor (erythroid-derived 2)-like 3 | Contains protein domain (PF00170) - bZIP transcription factor | transcriptfactor | 264488, 22278996, 35696286, 264091, 284259, 29331824, 29331826, 35696052, 264511, 55812038, 85658542, 264766, 21906765, 35695917, 264629, 35696423, 18108383, 87168518 |
| 2675 | 79563835 (5349, 5350) | | | UNCLASSIFIED | 264691 |
| 2676 | 79628393 (5351, 5352) | | | UNCLASSIFIED | 264906, 265008 |
| 2677 | 94329800 (5353, 5354) | Novel Protein sim. GBank gi 1079042 pir J52154 - acetyl-CoA synthetase - fruit fly (Drosophila melanogaster) | Contains protein domain (PF00501) - AMP-binding enzyme | synthase | 18108394, 65274572, 56182575, 22278994, 22278995, 56994075, 22278996, 35696286, 22278997, 22278999, 264490, 264259, 52645080, 29331824, 29331825, 60432289, 29331827, 29331828, 35696052, 29146499, 29331830, 284908, 52644045, 265006, 265007, 265008, 265009, 60432229, 60433356, 60433438, 55812038, 265010, 265011, 87168559, 265017, 265019, 18108351, 264682, 264448, 264683, 264288, 21906765, 21906766, 21906767, 21906768, 21906769, 55811957, 265020, 265021, 60170615, 52644150, 264691, 33657023, 263967, 33657109, 27486264, 27486285, 33657349, 35695763, 18108370, 18108374, 18108377, 55811576, 35696423, 35695855, 83373044, 18108387, 22279000, 22278002, 264564 |

| | | | | | |
|------|-----------------------|--|---|-------------------|---|
| 2678 | 95001694 (5355, 5356) | Novel Protein sim. GBank gij86760 pir A0465 - alanine transaminase (EC 2.6.1.2), cytosolic - human | | UNCLASSIFIED | 264488, 263994, 264489, 18108394, 52646842, 35696286, 22278999, 264259, 29331825, 35698052, 264508, 264905, 264509, 264906, 264907, 264908, 264909, 264510, 264511, 265006, 264512, 265007, 265008, 264910, 265009, 264593, 264594, 264596, 264758, 33109954, 52646317, 21808754, 265010, 265011, 87168559, 264800, 264601, 264602, 264603, 265017, 265019, 264605, 264760, 264762, 18108351, 264763, 264682, 264683, 264764, 264288, 264766, 264687, 264768, 264769, 264689, 21906765, 21906767, 35695917, 265020, 265021, 52644150, 264691, 33657023, 264693, 18108364, 18108365, 33657109, 33857349, 264628, 18108374, 35696423, 35695855, 264630, 264631, 264632, 264634, 264635, 264555, 264636, 264637, 264638, 264639, 18108385, 264563, 264564, 264565, 264566, 264486, 264567 |
| 2679 | 95361544 (5357, 5358) | Novel Protein sim. GBank gij1709233 sp P07514 NC5R_BOVIN - NADH- CYTOCHROME B5 REDUCTASE | Contains protein domain (PF00970) - FAD/NAD-binding Cytochrome reductase | UNCLASSIFIED | 264488, 22278996, 35696286, 264259, 29331826, 29331827, 29331828, 264909, 561182435, 264113, 264511, 265008, 60433438, 264758, 85658542, 87168474, 265011, 265017, 265019, 264288, 21906766, 21906767, 21906768, 21906769, 55811957, 265021, 264690, 33657023, 55810764, 35696423, 55811576, 264631, 18108381, 60170394, 83373044, 87168518, 264566 |
| 2680 | 87800356 (5359, 5360) | Novel Protein sim. GBank gij4589604 dbj BAA76824.1 - (AB023197) KIAA0980 protein [Homo sapiens] | Contains protein domain (PF00036) - EF hand | UNCLASSIFIED | 264259, 264102, 264905, 264908, 285007, 265008, 33109954, 265011, 18108351, 264768, 33657023, 20281149, 263972, 264630, 264635, 264638 |
| 2681 | 80933844 (5361, 5362) | Novel Protein sim. GBank gij1728837 sp P39194 ALU7_HUMAN - IIII ALU SUBFAMILY SQ WARNING ENTRY IIII | kinase | UNCLASSIFIED | 264489, 22278995, 264509, 264905, 264906, 264907, 264908, 264909, 264510, 265008, 264512, 264910, 264594, 60433438, 264758, 264803, 264604, 265019, 264605, 264760, 264764, 264687, 264768, 264769, 21906769, 35695917, 265020, 33657023, 264631, 264635, 264637, 264638, 264639, 264486 |
| 2682 | 94138934 (5363, 5364) | Novel Protein sim. GBank gij423468 pir JQ1974 - HTF9-C protein - mouse | Contains protein domain (PF00076) - RNA recognition motif. (a.k.a. RRM, RBD, or RNP domain) | UNCLASSIFIED | 35696286, 264908, 55811386, 265017, 55811150, 55811957, 35695917, 60431528, 55810764, 55811576, 35696423, 65274791, 56526486 |
| 2683 | 87774405 (5365, 5366) | Novel Protein sim. GBank gij5114351 gb AAD40286.1 - (AF156271) RING finger protein terf [Homo sapiens] | Contains protein domain (PF00822) - SPRY domain | interleukinrecept | 264909, 264769, 264635, 264636 |

| | | | | | |
|------|-----------------------|---|--|------------------|--|
| 2684 | 85787151 (5367, 5368) | Novel Protein sim. GBank gl 4886469 emb CAB43385.1 - (AL050284) hypothetical protein [Homo sapiens] | | | 264593 |
| 2685 | 88054209 (5369, 5370) | Novel Protein sim. GBank gl 3342729 (AC005331) - R31341_2 [Homo sapiens] | UNCLASSIFIED | | |
| 2686 | 87628690 (5371, 5372) | Novel Protein sim. GBank gl 4650844 db JBAA77027.1 - (AB026180) Kelch motif containing protein [Homo sapiens] | Contains protein domain (PF01344) - Kelch motif | struct | 264259, 29331822, 60432289, 29331827, 29331830, 264909, 264512, 264596, 264769, 264534, 264555, 264556, 264557, 264558, 60170394, 264559, 264486 |
| 2687 | 87998183 (5373, 5374) | Novel Protein sim. GBank gl 5281314 gb AAD41475.1 AF13312 - (AF133123) transcription factor IIC102 [Homo sapiens] | Contains protein domain (PF00515) - TPR Domain | transcriptfactor | 18108394, 18108396, 22278996, 35696286, 22278997, 29331826, 29331828, 66712502, 21906754, 265011, 264760, 264761, 264763, 264689, 21906765, 35696423, 264559, 18108385, 264563 |
| 2688 | 79959584 (5375, 5376) | Novel Protein sim. GBank gl 3880023 emb CAA97339 - (Z73098) Similarity to yeast hypothetical protein (Swiss Prot accession number Q09695); cDNA EST EMBL:D72982 comes from this gene; cDNA EST EMBL:D75947 comes from this gene [Caenorhabditis elegans] | Contains protein domain (PF00560) - Leucine Rich Repeat | | 264908, 264760 |
| 2689 | 94122440 (5377, 5378) | Novel Protein sim. GBank gl 4107276 emb CAA67130 - (X98506) acetyl-CoA synthetase [Solanum tuberosum] | Contains protein domain (PF00620) - RhoGAP domain | struct | 22278997, 22278998, 22278999, 29331824, 35696052, 264906, 264908, 56182435, 264512, 264910, 265009, 60433438, 21906754, 18108351, 264682, 264683, 264767, 21906765, 21906766, 21906768, 33657023, 33657182, 27486262, 27486264, 33657349, 18108370, 18108372, 18108374, 35696423, 35695855, 18108385, 22279002 |
| 2690 | 88003055 (5379, 5380) | Novel Protein sim. GBank gl 2477513 (AC002398) - F25965_3 [Homo sapiens] | | | |
| 2691 | 91218241 (5381, 5382) | Novel Protein sim. GBank gl 4107276 emb CAA67130 - (X98506) acetyl-CoA synthetase [Solanum tuberosum] | | synthase | 65274572, 56182575, 264259, 29331822, 29331824, 29331826, 29331828, 264112, 265009, 55812038, 264596, 33109954, 265017, 264448, 264288, 264369, 264684, 21906769, 60170615, 60431528, 55810764, 264634, 264636, 264556, 264637, 22279002, 264564, 264566 |
| 2692 | 94111914 (5383, 5384) | Novel Protein sim. GBank gl 3513303 (AC005594) - R26984_1 [Homo sapiens] | Contains protein domain (PF00326) - Prolyl oligopeptidase family | peptidase | |
| 2693 | 20438807 (5385, 5386) | Novel Protein sim. GBank gl 3122400 sp O35682 MJUG_MOUSE - MYELOID UPREGULATED PROTEIN | UNCLASSIFIED | | 264592 |
| 2694 | 94111918 (5387, 5388) | Novel Protein sim. GBank gl 4972740 gb AAD34765.1 - (AF132177) unknown [Drosophila melanogaster] | | | 264559 |
| 2695 | 95345513 (5389, 5390) | Novel Protein sim. GBank gl 4972740 gb AAD34765.1 - (AF132177) unknown [Drosophila melanogaster] | | collagen | 35696286, 56994075, 22278999, 264259, 35696052, 29331830, 265011, 264288, 56181562, 264690, 264692, 33657023, 27486262, 263976, 18108376, 35696423, 35695855, 60170394, 83373044, 56526486, 22279000, 22279002, 264566 |
| 2696 | 87874040 (5391, 5392) | Novel Protein sim. GBank gl 728831 sp P39188 ALU1_HUMAN - IIII ALU SUBFAMILY J WARNING ENTRY IIII | | synthase | 264594, 21906768, 18108370, 18108372 |

| | | | | | |
|------|-----------------------|---|---|-------------------|--|
| 2697 | 91638472 (5393, 5394) | Novel Protein sim. GBank gi 5689473 dbj BAA83020.1 - (ABD28991) KIAA1068 protein [Homo sapiens] | | UNCLASSIFIED | 35696286, 264259, 29331824, 29331825, 29331826, 29331827, 29331828, 35696052, 265006, 265007, 264512, 265009, 60170831, 60433356, 264595, 264758, 87168474, 265010, 265011, 87168559, 264601, 265017, 265018, 265019, 264781, 18108351, 284448, 264288, 264766, 264688, 264689, 21906765, 21906768, 265020, 265021, 60170615, 33657109, 18108376, 35696423, 35695855, 264555, 264558, 60170394, 264559, 18108387, 56526486, 87168518, 22279002, 264593, 264482 |
| 2698 | 94325691 (5395, 5396) | Novel Protein sim. GBank gi 841318 (U22818) - mutant sterol regulatory element binding protein-2 [Cricetulus griseus] | Contains protein domain (PF00412) - LIM domain containing proteins | transcript factor | 22278995, 35696286, 22278996, 22278997, 22278998, 22278999, 264259, 29331822, 29331824, 29331825, 60432289, 29331826, 29331827, 33656970, 264906, 28331830, 264909, 5264045, 264910, 60433356, 33657402, 33109954, 265017, 265018, 265019, 264288, 21906765, 21906766, 21906767, 21906769, 29148628, 35695917, 265021, 265022, 52844150, 33657023, 33657182, 27486261, 35696423, 65274781, 264638, 60432113, 22279000 |
| 2699 | 87780650 (5397, 5398) | Novel Protein sim. GBank | | UNCLASSIFIED | 264768, 18108357, 264690, 264691 |
| 2700 | 94139836 (5399, 5400) | Novel Protein sim. GBank gi 5174395 ref NP_005006.1 pB120 - Brain protein 120 | | UNCLASSIFIED | 60424269, 56182435, 60432228, 60433438, 55811386, 265017, 55811150, 264448, 56181562, 55811957, 264693, 33657109, 60431528, 264629, 55810764, 55811576, 65274791, 60431850, 60432113 |
| 2701 | 94148584 (5401, 5402) | Novel Protein sim. GBank gi 1572801 (U70854) - F38A5.1 gene product [Caenorhabditis elegans] | | | 18108394, 52645156, 35696286, 264259, 29331822, 29331824, 29331825, 60432289, 29331826, 29331827, 35696052, 29331828, 29148499, 265008, 60433356, 33657402, 60433438, 264595, 33657084, 18108351, 264288, 264769, 18108359, 21906768, 35695917, 33657023, 27486261, 18108374, 18108379, 35696423, 18108382, 83373044, 18108384, 18108388, 60432113, 22279000 |
| 2702 | 57295366 (5403, 5404) | Novel Protein sim. GBank gi 2605967 (AF030027) - 24 [Equine herpesvirus 4] | | UNCLASSIFIED | 29331828, 264512, 264555, 264556, 264557, 264558, 264559 |
| 2703 | 87648514 (5405, 5406) | Novel Protein sim. GBank gi 5689399 dbj BAA82983.1 - (AB028954) KIAA1031 protein [Homo sapiens] | Contains protein domain (PF00642) - Zinc finger C-x8-C-x5-C-x3-H type (and similar) | | 60432289, 265007, 21906765, 21906768, 265021, 264563 |

| | | | | | |
|------|-----------------------|---|--------------|--|--|
| 2704 | 87648515 (5407, 5408) | Novel Protein sim. GBank gij4335694[jgb/AAB63294] - (AF008554) Implantation-associated protein [Rattus norvegicus] | | | 264488, 22278995, 22278998, 29331828, 29146489, 264905, 264906, 264907, 52644045, 264511, 33857402, 264600, 264602, 265017, 264605, 264781, 18108351, 264764, 264687, 264769, 265021, 264691, 264692, 18108362, 264693, 18108370, 18108374, 264634, 264635 |
| 2705 | 87771745 (5409, 5410) | | | | 264489, 264509, 264511, 264512, 264910, 264593, 87168474, 264604, 264288, 264687, 264769, 264638, 264566, 264486 |
| 2706 | 94326789 (5411, 5412) | Novel Protein sim. GBank gij3255952[embjCAA16821.1] - (AL021728) /prediction=(method:; /match=(desc: [Drosophila melanogaster]) | UNCLASSIFIED | | 264488, 52646842, 65274572, 22278994, 56994075, 22278997, 264259, 29331824, 29331825, 29331826, 29331828, 33856970, 264907, 264908, 264909, 52644045, 56182435, 265006, 265007, 60433438, 55812038, 21906754, 52644298, 265010, 264601, 265017, 265019, 264681, 264448, 264682, 264288, 264686, 264687, 264688, 21906766, 21906769, 55811957, 35695917, 265020, 265021, 60170815, 264690, 264691, 33657023, 264692, 264693, 65274620, 27486264, 263972, 18108374, 18108377, 264635, 264636, 264556, 60170394, 83373044, 65274727, 87168518, 22278900 |
| 2707 | 88089839 (5413, 5414) | Novel Protein sim. GBank gij3417294 (AC004381) - Unknown gene product [Homo sapiens] | | | 22278996, 22278998, 56182435, 21906754, 87168559, 265017, 264448, 52645129 |
| 2708 | 91011351 (5415, 5416) | Novel Protein sim. GBank gij545790[bsj]147178 - DARPP-32-dopamine and cAMP-regulated phosphoprotein [human, brain, Peptide, 204 aa] | UNCLASSIFIED | | 65274572, 264259, 29331822, 29331825, 60432289, 29331828, 29331827, 29331828, 264909, 264510, 265007, 264910, 60433356, 60433438, 33109954, 265010, 265011, 264369, 264288, 264765, 264693, 264565 |
| 2709 | 94853988 (5417, 5418) | Novel Protein sim. GBank gij3169705 (AC004780) - F17127, 1 [Homo sapiens] | UNCLASSIFIED | | 29331822, 18108370, 18108374, 83373044 |
| 2710 | 87627979 (5419, 5420) | Novel Protein sim. GBank gij4468311[embjCAB37992] - (AL031432) dJ465N24.2.1 (PUTATIVE novel protein) (isoform 1) [Homo sapiens] | | | 29331824, 264759, 264693, 18108382, 18108388 |

| | | | | | |
|------|-----------------------|---|---|--------------|---|
| 2711 | 94111920 (5421, 5422) | Novel Protein sim. GBank gij3122400 sp O35682 MUG_MOUSE - MYELOID UPREGULATED PROTEIN | | UNCLASSIFIED | 264488, 264687, 52645156, 264769, 21908764, 21908765, 21906767, 21906768, 21906769, 55811957, 56994075, 22278997, 22278998, 265020, 265021, 264690, 264259, 264691, 264692, 33657023, 29331822, 29331824, 60424269, 29331826, 33657182, 29331827, 27486262, 33657349, 264508, 264905, 264907, 60431528, 264908, 264909, 55810764, 35696423, 65274781, 35695855, 265007, 264910, 60431850, 60432229, 264557, 264558, 55812038, 33108954, 18108385, 21906754, 33657084, 87168518, 87168474, 87168559, 60432113, 265017, 22279000, 265018, 265019, 22279002, 264760, 55811150, 264681, 18108351, 264585, 264764, 264568, 264288, 264768 |
| 2712 | 94312071 (5423, 5424) | Novel Protein sim. GBank gij5081315 gb AAD39343.1 AF076607 prediabetic NOD sera-reactive autoantigen [Mus musculus] | Contains protein domain (PF00515) - TPR Domain | transferase | 264488, 35696288, 22278998, 264259, 29331824, 60432289, 35696052, 264508, 264908, 66712502, 52644045, 265008, 60432229, 33657402, 60433358, 265010, 265019, 18108351, 264681, 264288, 264685, 21906765, 21906766, 21906768, 21906769, 55811957, 35695917, 265020, 265021, 60170815, 264691, 264692, 33657023, 33657109, 33657182, 33657349, 18108370, 18108374, 35696423, 35695855, 264555, 52644332, 56182323, 87168518, 60432113 66714117, 264906, 264563 |
| 2713 | 88003064 (5425, 5426) | Novel Protein sim. GBank gij2477513 (AC002398) - F25965_3 [Homo sapiens] | | UNCLASSIFIED | |
| 2714 | 13528218 (5427, 5428) | | | UNCLASSIFIED | 264636 |
| 2715 | 94122454 (5429, 5430) | Novel Protein sim. GBank gij4321968 gb AAD15897 - (AF067430) Smarce1-related protein [Mus musculus] | | UNCLASSIFIED | 264508, 264905, 264907, 264908, 264909, 264910, 264758, 265011, 264760, 264764, 264288, 264766, 264768, 264636 |
| 2716 | 88003068 (5431, 5432) | Novel Protein sim. GBank gij2477513 (AC002398) - F25965_3 [Homo sapiens] | | glycoprotein | 264091, 264259, 29331822, 66714117, 264908, 264369, 264693, 264556, 264563 264593, 264558 |
| 2717 | 80077461 (5433, 5434) | Novel Protein sim. GBank gij3327046 db BAA31591 - (AB014516) KIAA0616 protein [Homo sapiens] | | | |
| 2718 | 78604062 (5435, 5436) | | | | 264693 |
| 2719 | 88180423 (5437, 5438) | Novel Protein sim. GBank gij746495 (U23515) - weakly similar to gastrula zinc finger protein [Caenorhabditis elegans] | | UNCLASSIFIED | 29331822, 87168559, 265019, 265021, 52644150, 264691 |

| | | | | | |
|------|-----------------------|--|--|--------------|--|
| 2720 | 95086242 (5439, 5440) | Novel Protein sim. GBank gij1335873 (U46680) - ATP-dependent RNA helicase [Mus musculus] | Contains protein domain (PF00270) - DEAD/DEAH box helicase | helicase | 18108374, 60424179, 264489, 56182435, 21908765, 21908768, 35698423, 22278997, 265020, 265022, 265006, 265008, 264082, 264638, 60432229, 284691, 264692, 33657023, 264693, 33657402, 83373044, 29331824, 18108368, 60424269, 29331826, 18108385, 52645129, 21908754, 35696052, 29331828, 87168474, 264100, 265010, 265011, 265019, 22279002, 264905, 264482, 264563, 264906, 18108351, 264681, 18108370, 29331830, 264908, 68712502, 52644045, 264909, 264828, 18108354, 22278995, 35696286, 264259, 29331822, 29331824, 66714117, 29331826, 264906, 60433438, 265017, 18108351, 264448, 264288, 264769, 21908768, 265021, 33657109, 263969, 60431528, 264629, 55811576, 65274791, 35695855, 264631, 264637, 60170394, 56182323, 22279000 |
| 2721 | 95345523 (5441, 5442) | Novel Protein sim. GBank gij4929683 [gb]AAD34092.1 [AF15185 - (AF151855) CGI-97 protein [Homo sapiens] | Contains protein domain (PF01172) - Uncharacterized protein family UPF0023 | | 22278995, 35696286, 264259, 29331822, 29331824, 66714117, 29331826, 264906, 60433438, 265017, 18108351, 264448, 264288, 264769, 21908768, 265021, 33657109, 263969, 60431528, 264629, 55811576, 65274791, 35695855, 264631, 264637, 60170394, 56182323, 22279000 |
| 2722 | 91638807 (5443, 5444) | Novel Protein sim. GBank gij3212997 [gb]AAC23434.1 - (AC004997) match to ESTs AA667999 (NID:g2626700), AA165465 (NID:g1741481), Z45871 (NID:g575105), and TB4026 (NID:g712314); similar to various tre-like proteins including: AF040654 (PID:g2746883), D13644 (PID:g2104571), AL02114... | Contains protein domain (PF00566) - TBC domain | oncogene | 35696286, 22278999, 21906754, 265017, 284762, 264288, 21908765, 21908767, 21906768, 35695917, 18108362, 27486262, 35695855, 284558, 264559 |
| 2723 | 87387732 (5445, 5446) | | | UNCLASSIFIED | 264508, 264509, 264906, 264909, 264910, 55812038, 264766, 264687, 264629, 264636, 264486 |
| 2724 | 87639563 (5447, 5448) | Novel Protein sim. GBank gij4680681 [gb]AAD27730.1 [AF13295 - (AF132955) CGI-21 protein [Homo sapiens] | | ubiquitin | 18108396, 22278999, 20281099, 29331824, 29331826, 60432289, 29331828, 60170831, 60432229, 60433438, 18108351, 264682, 21908768, 21908767, 21908769, 35695917, 33657023, 33657109, 18108372, 18108374, 35695855, 22279000, 22279002 |
| 2725 | 94853991 (5449, 5450) | Novel Protein sim. GBank gij3169705 (AC004780) - F17127_1 [Homo sapiens] | | UNCLASSIFIED | 264488, 52644507, 264259, 29331827, 21906754, 265011, 18108351, 264448, 264288, 264685, 264689, 35695917, 265020, 33657182, 27486261, 18108370, 18108374, 35696423, 18108385, 22279000 |
| 2726 | 86880599 (5451, 5452) | Novel Protein sim. GBank gij3342738 (AC005328) - R26680_1, partial CDS [Homo sapiens] | | MHC | 264488, 264828, 264685 |

| | | | | | |
|------|-----------------------|---|---|----------------------|---|
| 2727 | 91010470 (5453, 5454) | Novel Protein sim. GBank gi 731267 sp P39219 RLUA_ECOLI - RIBOSOMAL LARGE SUBUNIT PSEUDOURIDINE SYNTHASE A (PSEUDOURIDYLATE SYNTHASE) (URACIL HYDROLYASE) | Contains protein domain (PF00849) - RNA pseudouridylylate synthase | synthase | 65274572, 56182575, 22278984, 56994075, 22278986, 22278987, 22278989, 60432049, 29331822, 29331824, 66714117, 29331826, 29331827, 35696052, 29331828, 33656970, 264509, 66712502, 264910, 33657402, 60433438, 264758, 55812038, 21906754, 33657084, 55811386, 265018, 265019, 264767, 21906765, 21906767, 21906769, 55811957, 35695917, 52644150, 33657023, 33657109, 33657182, 27486261, 27486262, 27486265, 33657349, 55811576, 35696423, 35695855, 264630, 60431850, 264636, 56182323, 87168518, 60432113, 22279000, 264564, 264565 |
| 2728 | 94126022 (5455, 5456) | Novel Protein sim. GBank gi 3880433 emb CAA91399 - (Z66521) similar to mitochondrial RNA splicing MSR4 like protein; cDNA EST EMBL:C09217 comes from this gene [Caenorhabditis elegans] | Contains protein domain (PF00153) - Mitochondrial carrier proteins | UNCLASSIFIED | 18108394, 56182435, 21906767, 55811957, 35695955, 265021, 264690, 264556, 264259, 264557, 29331822, 264559, 264448, 264288 |
| 2729 | 94126024 (5457, 5458) | Novel Protein sim. GBank gi 2408095 emb CAB16300 - (Z99168) putative RNA splicing protein [Schizosaccharomyces pombe] | Contains protein domain (PF00153) - Mitochondrial carrier proteins | UNCLASSIFIED | 65274572, 264259, 60432289, 66712502, 56182435, 264448, 264288, 264369, 55811957, 265021, 264557, 60432113 |
| 2730 | 94126026 (5459, 5460) | Novel Protein sim. GBank gi 3880433 emb CAA91399 - (Z66521) similar to mitochondrial RNA splicing MSR4 like protein; cDNA EST EMBL:C09217 comes from this gene [Caenorhabditis elegans] | Contains protein domain (PF00153) - Mitochondrial carrier proteins | transport | 264887, 29331824, 29331826, 35696052, 264107, 56182435, 265008, 265009, 264592, 60431735, 265011, 264601, 265017, 18108351, 264288, 29148627, 55811957, 265021, 264690, 18108368, 18108374, 264557, 264558, 264559, 18108387, 56528486, 264566, 264486 |
| 2731 | 87723022 (5461, 5462) | Novel Protein sim. GBank gi 1723239 sp Q10166 YAUB_SCHPO - HYPOTHETICAL 35.7 KD PROTEIN C26A3.11 IN CHROMOSOME 1 | Contains protein domain (PF00795) - Carbon-nitrogen hydrolase | | 264259, 35696052, 265008, 264758, 284762, 264448, 264288, 29148627, 21906769, 87168518, 22279002 |
| 2732 | 94126028 (5463, 5464) | Novel Protein sim. GBank gi 3880433 emb CAA91399 - (Z66521) similar to mitochondrial RNA splicing MSR4 like protein; cDNA EST EMBL:C09217 comes from this gene [Caenorhabditis elegans] | Contains protein domain (PF00153) - Mitochondrial carrier proteins | UNCLASSIFIED | 264687, 264489, 18108358, 56182435, 264689, 35696423, 55811957, 265021, 265006, 265008, 264910, 265009, 264690, 264555, 264259, 264556, 264557, 264558, 264559, 18108383, 33657109, 87168518, 265010, 264601, 60432113, 265017, 264905, 264448, 263972, 264369, 264567 29331825, 264509, 264909 |
| 2733 | 87363060 (5465, 5466) | | | UNCLASSIFIED | |
| 2734 | 94140286 (5467, 5468) | Novel Protein sim. GBank gi 4519621 db BAA75670.1 - (AB017614) OASIS protein [Mus musculus] | Contains protein domain (PF00170) - bZIP transcription factor | transcription factor | 60424179, 52644507, 56182575, 264259, 29331828, 264907, 264510, 264910, 60433356, 265019, 55811150, 264681, 264763, 264687, 33657182, 18108370, 60431528, 60431850, 56182323, 83373044 |

| | | | | | |
|------|-----------------------|--|--|--------------|--|
| 2735 | 87712338 (5469, 5470) | Novel Protein sim. GBank gij3850569 (AC005278) - ESTs gb T21276, gb T45403, and gb AA586113 come from this gene. [Arabidopsis thaliana] | | glycoprotein | 22278996, 60432289, 29331827, 29146498, 264108, 264909, 264112, 33657402, 87168474, 265017, 264762, 264448, 264764, 264684, 21908765, 264693, 33657109, 263976, 264636, 264638, 264557, 22279000, 22279002, 264567 |
| 2736 | 80247655 (5471, 5472) | | | UNCLASSIFIED | 264905, 264628, 264629, 263978, 264632, 264564 |
| 2737 | 87604526 (5473, 5474) | | | | 264680 |
| 2738 | 85731808 (5475, 5476) | Novel Protein sim. GBank gij2558501 dbj BAA22886 (D63850) hepatoma-derived growth factor [Mus musculus] | | | 264488, 265009, 264768, 264691 |
| 2739 | 94319834 (5477, 5478) | Novel Protein sim. GBank gij5420387 emb CAB46679.1 - (AJ243459) proteophosphoglycan [Leishmania major] | | UNCLASSIFIED | 264684, 83373044, 264566 |
| 2740 | 94148762 (5479, 5480) | Novel Protein sim. GBank gij3417386 emb CAA75495 - (Y15197) microtubule-associated protein, MAP-115 [Mus musculus] | | UNCLASSIFIED | 264488, 56182575, 22278995, 35696286, 22278997, 22278998, 22278999, 264259, 29331822, 29331824, 29331825, 29331827, 35696052, 29331828, 29146498, 29331830, 265006, 265007, 265009, 60432229, 33657402, 55812038, 87168474, 265010, 265011, 265017, 265018, 265019, 264605, 264681, 264288, 264369, 52644229, 21908765, 21906766, 21906767, 21906768, 21906769, 265020, 265022, 264691, 264692, 33657109, 18108370, 18108374, 55810764, 35695855, 264634, 60431850, 264639, 56182323, 18108382, 18108385, 65274727, 22279002, 264584 |
| 2741 | 88047518 (5481, 5482) | Novel Protein sim. GBank gij3242764 (AC005154) - similar to protein U28928 (PID:g861306) [Homo sapiens] | | UNCLASSIFIED | 22278996, 52644045, 52644229, 21906768, 21906769, 265020, 60170615, 264691 |
| 2742 | 87648644 (5483, 5484) | Novel Protein sim. GBank gij4758412 ref NP_004472.1 pGALN - UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetyltransferase 2 (GalNAc-T2) | Contains protein domain (PF00652) - Similarity to lectin domain of ricin beta-chain, 3 copies. | transferase | 264259, 264905, 264758, 55812038, 264369, 29148627 |
| 2743 | 87627991 (5485, 5486) | Novel Protein sim. GBank gij4468311 emb CAB37992 - (AL031432) cDNA EST EMBL:CO9217 comes from this gene (isoform 1) [Homo sapiens] | | UNCLASSIFIED | 35696286, 264259, 264906, 264908, 265006, 60433438, 265017, 18108351, 264448, 264764, 264288, 21906765, 21906767, 264690, 264691, 264693, 263969, 263971, 35695855, 264637, 264558, 18108382, 60432113 |
| 2744 | 94126030 (5487, 5488) | Novel Protein sim. GBank gij3880433 emb CAA91399 - (Z68521) similar to mitochondrial RNA splicing MSR4 like protein: cDNA EST EMBL:CO9217 comes from this gene [Caenorhabditis elegans] | Contains protein domain (PF00153) - Mitochondrial carrier proteins | kinase | 18108374, 264488, 56182435, 21906765, 35696423, 35695917, 35695855, 265020, 265021, 264511, 265009, 264490, 264558, 264259, 264557, 56182323, 264558, 264559, 18108383, 29331824, 18108385, 33657109, 29331826, 21908754, 29331827, 29331828, 33657349, 87168518, 265018, 264905, 264482, 264448, 264486, 264369, 264288 |

| | | | | | |
|------|-----------------------|---|---|--------------|--|
| 2745 | 87740125 (5489, 5490) | Novel Protein sim. GBank gij4405795[gb AAD19826] - (AF038963) RNA helicase [Homo sapiens] | Contains protein domain (PF00271) - Helicases conserved C-terminal domain | helicase | 35696286, 284509, 264905, 264907, 284908, 264510, 264512, 265008, 264758, 284801, 285017, 284804, 284763, 284288, 284886, 284769, 284693, 35696423, 35695855, 264634, 264636, 264563, 264564, 264565 |
| 2746 | 95418601 (5491, 5492) | Novel Protein sim. GBank gij4758738[re NP_004680.1 pMTA1 - metastasis associated 1 | Contains protein domain (PF00320) - GATA zinc finger | UNCLASSIFIED | 22278986, 22278988, 22278989, 29331822, 29331826, 29331827, 35696052, 29331828, 264905, 264806, 264907, 264908, 264909, 52644045, 265006, 60170831, 264596, 55812038, 265018, 264883, 264288, 21906765, 21906767, 21906768, 21906769, 265020, 264890, 33657023, 284693, 33657109, 18108368, 18108374, 264558, 18108385, 22279000, 264563 |
| 2747 | 94112677 (5493, 5494) | Novel Protein sim. GBank gij4557803[re NP_000282.1 pNPC1 - Niemann-Pick disease, type C1 | | glycoprotein | 264569, 52644507, 18108394, 22278995, 35696286, 22278997, 22278999, 52645080, 29331824, 56182181, 29331826, 29331827, 35696052, 264907, 264908, 264909, 265009, 33109854, 55811386, 87168474, 265010, 87168559, 264603, 265019, 264760, 264886, 264768, 21908769, 35695917, 60170815, 264692, 33657023, 52645129, 27486264, 60431528, 18108374, 35696423, 35695855, 264556, 56182323, 18108385, 264482 |
| 2748 | 91214983 (5495, 5496) | Novel Protein sim. GBank gij4181272[emb CAA09984] - (AJ012295) apaG protein [Rhizobium etli] | Contains protein domain (PF00646) - F-box domain. | | 65274572, 29331828, 264112, 264511, 265019, 264760, 264767, 264768, 264769, 21908768, 21908769, 265020, 27486262, 56526486, 87168518, 22279000 |
| 2749 | 87346307 (5497, 5498) | | | | 264259, 264908, 264510, 265008, 265009, 264760, 264369, 264768, 264563 |
| 2750 | 87336344 (5499, 5500) | Novel Protein sim. GBank gij1872498 (U74297) - PIUS [Oryctolagus cuniculus] | | UNCLASSIFIED | 264488, 52644507, 18108396, 56994075, 264259, 29331825, 29331826, 29331827, 29331828, 264508, 265009, 264910, 264591, 264595, 33657084, 265011, 265019, 18108351, 264288, 264686, 264769, 264889, 55811957, 264693, 27486264, 18108370, 18108374, 264558, 18108385, 264482, 264563 |
| 2751 | 87057465 (5501, 5502) | | | UNCLASSIFIED | 29331822, 29331824, 265017, 33657023 |
| 2752 | 88062875 (5503, 5504) | Novel Protein sim. GBank gij3041859 (AC004534) - OG-2 homeodomain protein-like; similar to U65067 (PID:g1575526) [Homo sapiens] | Contains protein domain (PF00046) - Homeobox domain | homeobox | |

| | | | | | |
|------|-----------------------|--|---|------------------|--|
| 2753 | 94138972 (5505, 5508) | Novel Protein sim. GBank gjl3851648 (AF088301) - neural F-box protein NFB42 [Rattus norvegicus] | Contains protein domain (PF00646) - F-box domain. | UNCLASSIFIED | 56182575, 56994075, 22278988, 28331822, 29331824, 29331825, 29331826, 265007, 264593, 55812038, 33109954, 18108351, 264288, 56181582, 21906767, 21906768, 285021, 264693, 18108374, 65274791, 264632, 58182323, 22279002, 264563, 264587 |
| 2754 | 94115513 (5507, 5508) | Novel Protein sim. GBank gjl535428 (U13736) - calmodulin-like protein [Plasmodium falciparum] | Contains protein domain (PF00038) - EF hand | struct | 22278989, 66714117, 29331827, 35696052, 29331828, 264906, 264908, 264909, 265011, 265017, 265018, 265019, 264288, 21906765, 21906767, 265022, 33657023, 264693, 56182323, 18108382, 22279000 |
| 2755 | 88001472 (5508, 5510) | Novel Protein sim. GBank gjl2996653 (AC004510) - R30385_2 [Homo sapiens] | Contains protein domain (PF00098) - Zinc finger, C2H2 type | transcriptfactor | |
| 2756 | 11465908 (5511, 5512) | Novel Protein sim. GBank gjl1173539 (U30473) - putative src-like adapter protein; non-catalytic src-like adapter protein containing SH3 and SH2 domains; homolog of mouse SLAP; Method: conceptual translation supplied by author [Homo sapiens] | Contains protein domain (PF00017) - Src homology domain 2 | UNCLASSIFIED | 264594 |
| 2757 | 95361590 (5513, 5514) | Novel Protein sim. GBank gjl3608372 (AF053768) - brain specific cortactin-binding protein CBP90 [Rattus norvegicus] | | eph | 85658542, 264693 |
| 2758 | 79637846 (5515, 5516) | Novel Protein sim. GBank gjl2072200 (U94863) - p40 [Borna disease virus] | | struct | 264693 |
| 2759 | 91005312 (5517, 5518) | Novel Protein sim. GBank gjl4914573 (embjCAB43885.1) - (AL050390) hypothetical protein [Homo sapiens] | Contains protein domain (PF00023) - Ank repeat | UNCLASSIFIED | 65274572, 35696286, 66714117, 29331828, 264508, 56182435, 21906754, 55811957, 264629, 264636, 56182323, 22279002, 264908 |
| 2760 | 79824798 (5519, 5520) | Novel Protein sim. GBank gjl3138150 (AF050182) - PERIOD 3 [Mus musculus] | Contains protein domain (PF00989) - PAS domain | nuclease | 22278998, 29331822, 29331830, 265010, 265019, 264288, 21906765, 21906768, 21906769, 265020, 56182323, 22279002, 264563 |
| 2761 | 87639597 (5521, 5522) | Novel Protein sim. GBank gjl2905843 (AF045244) - ribitol kinase [Klebsiella pneumoniae] | | struct | 18108394, 22278998, 264906, 264909, 265006, 265007, 264757, 265010, 265011, 265017, 265019, 18108351, 264448, 264683, 264686, 264768, 265020, 265021, 265022, 264691, 18108362, 264693, 18108365, 33657109, 18108368, 18108370, 18108381, 18108382, 18108384, 18108388, 87168518 |
| 2762 | 87592699 (5523, 5524) | Novel Protein sim. GBank gjl3511122 (AF060503) - zinc finger protein [Homo sapiens] | Contains protein domain (PF01352) - KRAB box | transcriptfactor | 264369, 35696423 |
| 2763 | 87539968 (5525, 5526) | Novel Protein sim. GBank gjl2905843 (AF045244) - ribitol kinase [Klebsiella pneumoniae] | Contains protein domain (PF00370) - FGGY family of carbohydrate kinases | | 52646842, 264259, 29331822, 29331825, 29331826, 29331828, 33656970, 264905, 264907, 29331830, 265006, 265009, 21906754, 265019, 264448, 21906769, 27466262, 56182323, 56526486, 87168518, 264487 |
| 2764 | 94305140 (5527, 5528) | Novel Protein sim. GBank gjl2905843 (AF045244) - ribitol kinase [Klebsiella pneumoniae] | | | |

| | | | | | |
|------|-----------------------|---|--|---------------|--|
| 2765 | 94315105 (5528, 5530) | Novel Protein sim. GBank gij488672 emb CAA17688.2 - (AL02018) /prediction=(method:; /prediction=(method:; /match=(desc: [Drosophila melanogaster]) | | | 264488, 65274572, 22278995, 35696286, 22278998, 22278999, 264259, 29331822, 29331824, 35696052, 264508, 264509, 264905, 264908, 264907, 264908, 264909, 264512, 264910, 265009, 264592, 264595, 264758, 55812038, 33109954, 265010, 87168559, 264600, 265018, 264760, 264761, 264762, 264763, 264448, 264764, 264288, 264766, 264767, 264768, 56181562, 21906764, 21906765, 21906768, 35695917, 265021, 264691, 264692, 33657023, 33657109, 264628, 18108370, 264629, 18108374, 55811576, 35696423, 35695855, 264631, 264634, 264635, 264636, 264637, 264638, 264558, 264639, 83373044, 87168516, 22279000, 22279002, 264563, 264482 |
| 2766 | 94315109 (5531, 5532) | Novel Protein sim. GBank gij5441611 emb CAB46854.1 - (AJ388555) hypothetical protein [Canis familiaris] | Contains protein domain (PF00076) - RNA recognition motif (a.k.a. RRM, RBD, or RNP domain) | - dna_ma_bind | 264488, 65274572, 60432289, 264907, 264909, 264511, 264512, 60433358, 264288, 264685, 264689, 35695917, 265022, 264693, 264628, 65274791, 264635, 264555, 264556, 264557, 264638, 264558, 264559, 83373044, 60432113 |
| 2767 | 80204297 (5533, 5534) | Novel Protein sim. GBank gij1079451 pir A55463. - tropomodulin, skeletal muscle - chicken | | struct | 264112, 263974, 264558 |
| 2768 | 94322238 (5535, 5536) | Novel Protein sim. GBank gij5441322 emb CAB46721.1 - (AL031427) dJ167A19.1 (novel protein) [Homo sapiens] | | UNCLASSIFIED | 264488, 263994, 264489, 35696286, 22278998, 22278999, 264259, 29331822, 60432289, 29331826, 35696052, 264508, 264509, 264905, 264908, 264907, 29331830, 264908, 264909, 264510, 264511, 264512, 265008, 264910, 265009, 264591, 264592, 264593, 33857402, 264594, 264595, 264757, 264596, 265011, 264604, 265019, 264760, 264681, 18108351, 264764, 264288, 264766, 264686, 264768, 18108357, 264768, 264689, 21906766, 21906768, 21906769, 35695917, 265021, 265022, 33657023, 33657109, 264628, 264629, 18108374, 35696423, 35695855, 264630, 264631, 264632, 264634, 264635, 264636, 264555, 264638, 264639, 83373044, 264563, 264564, 264565, 264566, 264488, 264567 |

| | | | | | |
|------|-----------------------|--|--|---------------|--|
| 2769 | 95311088 (5537, 5539) | Novel Protein sim. GBank gi 5419859 emb CAB46375.1 - (AL098725) hypothetical protein [Homo sapiens] | | tubulin | 264488, 56182575, 264259, 29331822, 29331824, 66714117, 29331825, 60432289, 29331828, 264908, 29331830, 56182435, 264592, 33657402, 264448, 264369, 264288, 60170615, 264691, 33657023, 264692, 33657109, 18108374, 55811576, 264634, 264636, 56182323, 83373044, 60432113, 18108398, 22278995, 22278996, 22278999, 264105, 265006, 265019, 18108351, 264687, 21906765, 18108364, 264629, 18108374, 264631, 18108385, 18108388 |
| 2770 | 87730182 (5539, 5540) | Novel Protein sim. GBank gi 5701965 emb CAB52157.1 - (AL109736) WD repeat protein [Schizosaccharomyces pombe] | Contains protein domain (PF00400) - WD domain, G-beta repeat | kinase | 18108398, 22278995, 22278996, 22278999, 264105, 265006, 265019, 18108351, 264687, 21906765, 18108364, 264629, 18108374, 264631, 18108385, 18108388 |
| 2771 | 88084071 (5541, 5542) | Novel Protein sim. GBank gi 3083433 (AC004125) - Unknown gene product [Homo sapiens] | | UNCLASSIFIED | 264259, 29331822, 29331824, 29331825, 264369 |
| 2772 | 95357308 (5543, 5544) | Novel Protein sim. GBank gi 4885531 ref NP_005465.1 pVY C - histone deacetylase 5 | Contains protein domain (PF00850) - Histone deacetylase family | histone | 264259, 29331822, 29331824, 66714117, 60432289, 29331827, 264905, 265009, 264592, 55812038, 65274444, 264766, 21906769, 33657109, 263978, 264555, 264638, 264557, 83373044, 264563, 264584, 264486 |
| 2773 | 94138994 (5545, 5546) | Novel Protein sim. GBank gi 3288888 (AC005253) - R28445_1 [Homo sapiens] | | UNCLASSIFIED | 18108374, 264686, 264687, 263976, 56182435, 264689, 55810784, 21906766, 35696423, 55811576, 65274791, 56181686, 55811957, 35695855, 264110, 265021, 264112, 265022, 265006, 265008, 264092, 264094, 60431850, 264637, 264691, 264259, 264692, 263981, 264594, 60433356, 56182323, 264693, 264757, 56182181, 55812038, 29331825, 60424269, 18108385, 29331826, 29331827, 27486261, 29331828, 35696052, 55811386, 264107, 60432113, 265017, 55811150, 18108351, 264681, 264906, 18108370, 264484, 264682, 20281069, 264448, 66712502, 264683, 264764, 264288, 264684, 264768, 263974, 22278995, 35696286, 22278996, 22278999, 264259, 29331826, 60432289, 35696052, 264112, 33657402, 21906754, 87168559, 265017, 265018, 264288, 21906766, 21906767, 21906769, 35695917, 265020, 265021, 33657023, 33657109, 18108370, 263976, 35696423, 35695855, 87168518, 22279000, 264482 |
| 2774 | 87818908 (5547, 5548) | Novel Protein sim. GBank gi 465852 sp P34388 YLS3_CAEEL - HYPOTHETICAL 70.7 KD PROTEIN F09G8.3 IN CHROMOSOME III | Contains protein domain (PF00380) - Ribosomal protein S9/S16 | ribosomalprot | 22278995, 35696286, 22278996, 22278999, 264259, 29331826, 60432289, 35696052, 264112, 33657402, 21906754, 87168559, 265017, 265018, 264288, 21906766, 21906767, 21906769, 35695917, 265020, 265021, 33657023, 33657109, 18108370, 263976, 35696423, 35695855, 87168518, 22279000, 264482 |

| | | | | | |
|------|-----------------------|--|--|---------------|--|
| 2775 | 95307987 (5549, 5550) | Novel Protein sim. GBank gi 4688132 gb AAD27775.1 AF07704 - (AF077042) 30S ribosomal protein S7 homolog [Homo sapiens] | Contains protein domain (PF00177) - Ribosomal protein S7p/S5a | ribosomalprot | 264488, 22278995, 56994075, 22278998, 35698288, 22278998, 22278999, 264259, 29331822, 29331824, 29331825, 60432289, 29331827, 29331828, 35696052, 265007, 21906754, 265017, 265019, 264448, 264882, 264369, 264288, 18108354, 52644229, 264689, 21906765, 21906768, 21906767, 21906768, 21906769, 35695917, 265021, 265022, 60170615, 264691, 18108370, 35696423, 65274791, 35695855, 264634, 60431850, 60170394, 56182323, 264558, 18108388, 22279000, 264563, 264565, 264488, 264567 |
| 2776 | 87781557 (5551, 5552) | | | | 56182575, 22278998, 22278999, 264259, 29331822, 29331824, 264908, 29331830, 264510, 33657402, 21906754, 5581386, 265017, 265019, 264448, 264288, 21906765, 21906766, 21906767, 21906768, 21906769, 265020, 265021, 265022, 60170615, 55810764, 55811578, 264555, 56526486, 22279000 |
| 2777 | 79818729 (5553, 5554) | | | UNCLASSIFIED | 264907, 264766 |
| 2778 | 82112411 (5555, 5556) | | | UNCLASSIFIED | 264907, 264593, 264760, 264628 |
| 2779 | 87649729 (5557, 5558) | Novel Protein sim. GBank gi 4680711 gb AAD27745.1 AF13297 - (AF132970) CGI-36 protein [Homo sapiens] | | UNCLASSIFIED | 22278997, 264259, 29331824, 56714117, 35698052, 265006, 264512, 264448, 264288, 29148827, 18108364, 20281149, 18108370, 264629 |
| 2780 | 94679397 (5559, 5560) | Novel Protein sim. GBank gi 4758524 ref NP_004825.1 pH-GK] - HPK/GCK-like kinase | Contains protein domain (PF00780) - CNH domain | kinase | 29146499, 65274791, 264634, 264639 |
| 2781 | 91220057 (5561, 5562) | Novel Protein sim. GBank gi 4469352 gb AAD21222] - (AF069502) ubiquitin specific protease UBP43 [Mus musculus] | Contains protein domain (PF00443) - Ubiquitin carboxyl-terminal hydrolase family 2 | ubiquitin | 60424179, 29331824, 60424269, 66714117, 29331826, 56182435, 87168474, 265017, 264764, 56181562, 21906765, 21906766, 21906768, 35695917, 265020, 33657023, 35695855, 56182323, 87168518 |
| 2782 | 94233146 (5563, 5564) | Novel Protein sim. GBank gi 4505013 ref NP_002310.1 pLRN] - leucine-rich neuronal protein | Contains protein domain (PF00560) - Leucine Rich Repeat | struct | 65274572, 22278996, 22278998, 60432049, 264259, 29331822, 29331824, 29331826, 60432289, 29331828, 264905, 264907, 264908, 264909, 52844045, 265009, 265017, 265018, 264604, 265019, 264760, 264683, 264288, 264766, 264685, 264688, 264768, 52644229, 264689, 21906768, 265020, 265021, 264691, 18108362, 264692, 33657023, 264693, 33657109, 33657349, 18108370, 264628, 263978, 35695855, 264557, 56182323, 83373044, 18108385 |

| | | | | | |
|------|-----------------------|---|---|--------------|---|
| 2783 | 80016828 (5585, 5586) | Novel Protein sim. GBank gi 728831 sp P39188 ALU1_HUMAN - IIII ALU SUBFAMILY J WARNING ENTRY IIII | | tm7 | 264908, 264628, 263978, 263981 |
| 2784 | 87614360 (5567, 5568) | | | | 264259, 28331822, 28331824, 29331825, 264482 |
| 2785 | 88071930 (5569, 5570) | Novel Protein sim. GBank gi 2134933 pir S58890 - collapsin response mediator protein - human | | UNCLASSIFIED | 22278999, 264908, 264758, 265018, 264769, 21906765, 21906768, 21906769, 265020, 264584 |
| 2786 | 87408542 (5571, 5572) | Novel Protein sim. GBank gi 2073564 (U80223) - eukaryotic initiation factor eIF-2 alpha kinase; DGCN2 [Drosophila melanogaster] | Contains protein domain (PF00069) - Eukaryotic protein kinase domain | kinase | 264905, 265017 |
| 2787 | 87901266 (5573, 5574) | Novel Protein sim. GBank gi 5174507 ref NP_006020.1 pMA1 - paraneoplastic neuronal antigen | | UNCLASSIFIED | 264768, 21906765, 21906768, 55811957, 22278999, 264093, 60170815, 264259, 29331822, 18108365, 29331824, 33657109, 29331827, 35696052, 264100, 264105, 264908, 263977, 55811576, 264635, 264637, 60433438, 60432113, 265017, 265019, 22279002, 55811150, 264369, 264288 |
| 2788 | 88080844 (5575, 5576) | Novel Protein sim. GBank gi 3252826 (AC004382) - Unknown gene product [Homo sapiens] | | UNCLASSIFIED | |
| 2789 | 85491275 (5577, 5578) | Novel Protein sim. GBank gi 2495729 sp Q92556 Y281_HUMAN - HYPOTHETICAL PROTEIN KIAA0281 (HA6725) | | UNCLASSIFIED | 264103, 21906769, 264693 |
| 2790 | 87602784 (5578, 5580) | Novel Protein sim. GBank gi 5101772 emb CAB45135.1 - (AJ242978) p821 [Homo sapiens] | | strut | 264488, 264768, 264910, 264631, 264636, 264690, 264691, 264259, 264638, 29331824, 264508, 264509, 264905, 264563, 264908, 264628, 18108370, 264907, 264764, 264908, 264288, 264909 |
| 2791 | 88083195 (5581, 5582) | Novel Protein sim. GBank gi 2911266 (AC002550) - Unknown gene product [Homo sapiens] | | | 21906764, 18108368 |
| 2792 | 85083783 (5583, 5584) | Novel Protein sim. GBank gi 2854163 gb AAC02581.1 - (AF045642) No definition line found [Caenorhabditis elegans] | | | 22278996, 22278997, 264259, 29331822, 29331824, 29331825, 66714117, 29331826, 60432289, 29331827, 35696052, 264908, 66712502, 29331830, 264909, 60432229, 60433356, 60433438, 33109954, 265010, 265011, 265017, 265018, 265019, 264682, 264448, 264288, 264369, 264766, 52644229, 21906765, 21906768, 265020, 265021, 33657023, 263974, 18108374, 65274781, 35695855, 264636, 264556, 264558, 56182323, 83373044, 18108385, 56526486, 60432113, 22279000, 264567 |
| 2793 | 87425476 (5585, 5586) | | | UNCLASSIFIED | 264259, 60432289, 66712502, 265009, 264636 |
| 2794 | 85794830 (5587, 5588) | | | UNCLASSIFIED | 264689, 265022, 264691, 18108388, 264567 |

| | | | | | |
|------|-----------------------|---|---|--------------|--|
| 2785 | 95334888 (5589, 5590) | Novel Protein sim. GBank gij5454148jrefjNP_006348.1 pUBE2 - ubiquitin-conjugating enzyme E2E 3 (homologous to yeast UBCA15) | Contains protein domain (PF00179) - Ubiquitin-conjugating enzyme | ubiquitin | 65274572, 56182575, 35696286, 22278988, 22278999, 60432049, 264259, 29331822, 29331825, 66714117, 60432289, 29331828, 29331827, 35696052, 29331828, 264907, 66712502, 56182435, 264511, 265007, 264512, 264910, 60432229, 60433356, 60433438, 33109954, 85658542, 265018, 265019, 264288, 264686, 21906764, 21906765, 21906768, 21906768, 21906769, 55811957, 265020, 265021, 265022, 52644150, 33657023, 264693, 65274620, 33657109, 35696423, 55811576, 65274791, 56182323, 56526486, 60432113, 22279002, 264482, 264563, 264484, 264567 18108394, 65274572, 56182575, 56994075, 22278999, 264490, 60432049, 264259 29331822, 29331824, 29331826, 35696052, 264509, 264906, 264907, 264908, 66712502, 56182435, 264510, 265006, 264512, 265007, 265008, 264910, 265009, 264591, 264592, 60433356, 264594, 264595, 55812038 264596, 21906754, 60174639, 87168474, 265010, 265011, 265017, 265018, 265019, 55811150, 264762, 18108351, 264448, 264682, 264369, 264288, 264685, 264766, 264687, 56181562, 264769, 264689, 21906765, 21906766, 21906767, 21906768, 21906769, 55811957, 35695917, 265020, 265021, 60170815, 52844150, 264692, 33657023, 18108362, 264693, 65274620, 33657109, 33657182, 27486265, 33657349, 18108374, 35696423, 65274791, 35695855, 264556, 264557, 56182323, 264558, 60170394, 83373044, 65274727, 87168518, 22278000, 264563, 264564, 264585, 264566, 264567 |
| 2786 | 94848857 (5591, 5592) | Novel Protein sim. GBank gij4680851 gb AAD27715.1 AF13284 - (AF132940) CGI-06 protein [Homo sapiens] | | UNCLASSIFIED | |
| 2787 | 95110780 (5593, 5594) | Novel Protein sim. GBank gij4838557 gb AAD31040.1 - (AF143859) potassium channel modulatory factor DEBT-91 [Mus musculus] | Contains protein domain (PF00569) - Zinc finger present in dystrophin, CBP/p300 | UNCLASSIFIED | 56182575, 22278995, 22278996, 22278997, 29331827, 29148499, 264509, 264908, 56182435, 264757, 21906754, 265010, 265017, 265018, 264681, 264682, 264683, 264686, 21906765, 21906767, 21906768, 21906769, 29148629, 265020, 52844150, 264690, 33657182, 264629, 18108376, 56182323, 22279002, 264563 264488, 264490, 264259, 264448, 20281149, 20281152, 264556, 264557, 264558, 264559, 264483, 264486, 264567 |
| 2788 | 86198005 (5595, 5596) | Novel Protein sim. GBank gij2852645 (AF007160) - unknown [Homo sapiens] | | UNCLASSIFIED | |

| | | | | | |
|------|-----------------------|---|---|--------------|--|
| 2799 | 86090651 (5597, 5598) | Novel Protein sim. GBank gij3252825 (AC004382) - Unknown gene product [Homo sapiens] | | UNCLASSIFIED | 22278996, 22278997, 22278998, 22278999, 60432049, 264259, 29331826, 35696052, 29331828, 264909, 60433356, 33657402, 33109954, 87168474, 264448, 52644229, 21906766, 21906767, 21906768, 35695917, 265020, 265021, 265022, 52644150, 35695955, 264634, 60432113, 22279000 |
| 2800 | 88316481 (5599, 5600) | Novel Protein sim. GBank gij4240301[d]BAA74928.1] - (AB020713) KIAA0906 protein [Homo sapiens] | | glycoprotein | 264488, 56994075, 264259, 20281099, 29331825, 29331827, 264905, 56182435, 265006, 265011, 87168559, 265017, 265019, 264448, 264288, 264766, 264686, 60170615, 264691, 264692, 27486265, 264628, 264629, 264636, 264557, 264558, 264559, 87168518, 264564, 264566, 264567 |
| 2801 | 86068814 (5601, 5602) | | Contains protein domain (PF00627) - UBA domain | | 265007, 264687 |
| 2802 | 88082477 (5603, 5604) | Novel Protein sim. GBank gij2337865 (AC002464) - organic callon transporter; 50% similarity to JC4884 (PID:g2143892) [Homo sapiens] | Contains protein domain (PF00083) - Sugar (and other) transporter | transport | 264448, 35694855 |
| 2803 | 79577446 (5605, 5606) | Novel Protein sim. GBank | | UNCLASSIFIED | 264639 |
| 2804 | 57111131 (5607, 5608) | gij4559368[g]AAD23029.1 AC00658 - (AC006585) hypothetical protein [Arabidopsis thaliana] | Contains protein domain (PF01585) - G-patch domain | peptidase | 264566 |
| 2805 | 87398486 (5609, 5610) | | | | 264092, 264259, 29331822, 29331824, 29331827, 29331828, 264508, 265007, 265009, 21906754, 264369, 264288, 264632, 60170394, 264563, 264482 |
| 2806 | 87898951 (5611, 5612) | Novel Protein sim. GBank gij1168973 sp P444403 CLPB_HAEIN - CLPB PROTEIN | | UNCLASSIFIED | 22278995, 22278996, 22278999, 264259, 29331822, 29331824, 29331826, 29331827, 29146498, 265008, 265009, 60433438, 265017, 265018, 265019, 264448, 264288, 21906765, 21906767, 21906769, 29148629, 265022, 52644150, 56182323 |

| | | | | | |
|------|-----------------------|--|---|----------------------|--|
| 2807 | 91720702 (5613, 5614) | Novel Protein sim. GBank gi 4468310 emb CAB37991 - (AL031432) dJ465N24.1 (PUTATIVE novel protein similar to predicted yeast and worm proteins) [Homo sapiens] | | UNCLASSIFIED | 52644507, 52645156, 52646842, 18108398, 56182575, 22278994, 22278995, 22278996, 56994075, 35696286, 22278997, 22278998, 22278999, 264259, 52845080, 29331822, 29331824, 29331825, 29331826, 35696052, 29331828, 33656970, 264100, 264105, 264907, 52644045, 60433356, 264594, 60433438, 52646317, 21906754, 33109954, 33657084, 52644296, 87168474, 265011, 87168559, 265017, 265018, 265019, 18108351, 264448, 264763, 264687, 52644229, 21906765, 21906766, 21906767, 21906768, 21906769, 35695917, 265021, 265022, 52644150, 33657023, 52645129, 33657109, 33657182, 27486261, 27486262, 27486264, 27486265, 35695763, 18108370, 18108372, 18108374, 18108376, 263977, 18108377, 35698423, 35695855, 52644332, 83373044, 18108385, 18108387, 87168518, 60432113 |
| 2808 | 95359111 (5615, 5616) | Novel Protein sim. GBank gi 5541863 emb CAB51071.1 - (AL096857) hypothetical protein [Homo sapiens] | | MHC | 60432289, 264510, 265010, 265017, 265018, 264681, 264686, 265021, 264690, 22279000, 264566 |
| 2809 | 88083530 (5617, 5618) | Novel Protein sim. GBank gi 27272561 (AC004002) - similar to ciliary dynein beta heavy chain, 78% Similarity to P23098 (PID:gi118965) [Homo sapiens] | | ATPase_associated | 18108351 |
| 2810 | 87259032 (5619, 5620) | | | UNCLASSIFIED | 264569, 22278996, 264091, 264259, 29331828, 29146499, 29146629, 29146784, 264693 |
| 2811 | 91235845 (5621, 5622) | Novel Protein sim. GBank gi 3264583 (AC005189) - match to ESTs H97758 (NID:gi1118643) and AA085546 (NID:gi1628773) [Homo sapiens] | | UNCLASSIFIED | 22278999, 264259, 66712502, 264693 |
| 2812 | 88093334 (5623, 5624) | Novel Protein sim. GBank gi 4240273 dbj BAA74915.1 - (AB020699) KIAA0892 protein [Homo sapiens] | | UNCLASSIFIED | 264106 |
| 2813 | 91218765 (5625, 5626) | Novel Protein sim. GBank gi 3548791 (AC005620) - R33590_1 [Homo sapiens] | | UNCLASSIFIED | 56182575, 29331828, 264906, 66712502, 55811386, 265017, 265018, 265019, 264689, 265020, 87168518, 60432113 |
| 2814 | 90880906 (5627, 5628) | Novel Protein sim. GBank gi 3548791 (AC005620) - R33590_1 [Homo sapiens] | Contains protein domain (PF00170) - bZIP transcription factor | transcription factor | 55274572 |
| 2815 | 79774521 (5629, 5630) | Novel Protein sim. GBank gi 5420389 emb CAB46680.1 - (AJ243460) proteophosphoglycan [Leishmania major] | | UNCLASSIFIED | 264907, 264909 |
| 2816 | 95358229 (5631, 5632) | | | UNCLASSIFIED | 264488, 35696286, 29331825, 29331828, 264508, 264509, 264905, 264906, 264907, 264908, 264909, 264511, 264910, 265011, 264760, 264681, 264766, 264769, 264689, 21906765, 264693, 264628, 18108370, 264629, 264631, 264634, 264563, 264564, 264566, 264486 |

| | | | | | |
|------|-----------------------|--|--|--------------|--|
| 2817 | 87749542 (5633, 5634) | Novel Protein sim. GBank gi 1293846 (U56966) - coded for by C. elegans cDNA yk30b3.5; coded for by C. elegans cDNA yk30b3.3 [Caenorhabditis elegans] | | | 264259, 29331822, 29331827, 264508, 264509, 264905, 264907, 264908, 264909, 56182435, 264510, 265006, 264511, 264512, 264593, 264758, 265010, 264760, 264781, 264784, 264288, 264687, 264769, 55811957, 35695917, 33657109, 263978, 264634, 264636, 264639, 264584, 264565, 264566, 264486, 264567 |
| 2818 | 88073579 (5635, 5636) | Novel Protein sim. GBank gi 549986 (U13149) - possible apoptosis-associated protein [Pennisetum ciliare] | Contains protein domain (PF00023) - Ank repeat | UNCLASSIFIED | 66712502 |
| 2819 | 87793527 (5637, 5638) | Novel Protein sim. GBank gi 4929773 (gb AAD34147.1) (AF15209) (AF152097) CGI-05 protein [Homo sapiens] | Contains protein domain (PF00919) - Uncharacterized protein family UPF0004 | UNCLASSIFIED | 264909, 264511 |
| 2820 | 87765744 (5639, 5640) | | | UNCLASSIFIED | 18108394, 52646365, 52644045, 264112, 265009, 21908754, 265017, 18108351, 264683, 264369, 264686, 264769, 21906769, 33657023, 33657109, 18108370, 18108374, 18108377, 18108385 |
| 2821 | 95320511 (5641, 5642) | Novel Protein sim. GBank gi 399144 (sp P02747) (C1QC_HUMAN - COMPLEMENT C1Q SUBCOMPONENT, C CHAIN PRECURSOR) | Contains protein domain (PF00386) - C1q domain | complement | 264638 |
| 2822 | 94260221 (5643, 5644) | Novel Protein sim. GBank gi 2224671 (db BAA20820) (AB002363) KIAA0365 [Homo sapiens] | | UNCLASSIFIED | 264259, 29331822, 60432289, 264905, 60433356, 33657402, 265011, 264760, 21906765, 60170815, 264638 |
| 2823 | 85320513 (5645, 5646) | Novel Protein sim. GBank gi 399144 (sp P02747) (C1QC_HUMAN - COMPLEMENT C1Q SUBCOMPONENT, C CHAIN PRECURSOR) | Contains protein domain (PF00386) - C1q domain | complement | 264766 |
| 2824 | 95320515 (5647, 5648) | Novel Protein sim. GBank gi 399144 (sp P02747) (C1QC_HUMAN - COMPLEMENT C1Q SUBCOMPONENT, C CHAIN PRECURSOR) | | UNCLASSIFIED | 264907 |
| 2825 | 19742170 (5649, 5650) | Novel Protein sim. GBank gi 3859683 (emb CAA22020) (AL033503) conserved hypothetical protein [Candida albicans] | | | 264760 |
| 2826 | 94311905 (5651, 5652) | | | | 52645156, 22278994, 35696286, 22278997, 22278998, 52645080, 29331822, 29331824, 29331825, 29331827, 35696052, 29331828, 33656970, 52644045, 52646317, 33657084, 52644296, 265017, 265018, 265019, 264288, 21906764, 21906765, 21906766, 21906767, 21906769, 35695917, 52644150, 33657109, 33657182, 27486261, 27486262, 33657349, 27486265, 35695763, 35696423, 35695855, 52644332, 18108385, 87168518, 264484 |

| | | | | | |
|------|-----------------------|--|---|-------------------|--|
| 2827 | 85320519 (5653, 5654) | Novel Protein sim. GBank gi 398144 sp P02747 C1QC_HUMAN - COMPLEMENT C1Q SUBCOMPONENT, C CHAIN PRECURSOR | Contains protein domain (PF00386) - C1q domain | complement | 264488, 263994, 264489, 65274572, 29331822, 66714117, 29331827, 29331828, 264508, 264905, 264509, 264906, 264907, 264908, 264909, 264510, 265006, 264511, 265007, 264512, 265008, 264910, 264591, 60432229, 264592, 264593, 264595, 264596, 21908754, 265011, 264600, 264601, 264602, 265017, 264604, 264605, 264288, 264766, 264767, 264689, 55811957, 264534, 264690, 264691, 264692, 264693, 263972, 264629, 35895855, 264631, 264634, 264635, 264555, 264636, 264637, 264556, 264638, 264558, 83373044, 18108385, 60432113, 22279002, 264563, 264564, 264565, 264566, 264567, 264486, 18108391 |
| 2828 | 91228615 (5655, 5656) | Novel Protein sim. GBank gi 3598974 (AF077000) - protein tyrosine phosphatase TD14 [Rattus norvegicus] | Contains protein domain (PF00102) - Protein-tyrosine phosphatase | phosphatase | 29331822, 35898052, 264104, 264908, 265007, 264591, 265010, 265011, 265019, 264768, 264686, 55811957, 18108370, 18108374, 55810764, 35696423, 55811576, 56182323, 83373044, 87168518 22278996, 22278997, 264091, 264093, 60432049, 264259, 29331822, 29331825, 29331827, 29331828, 264905, 264509, 66712502, 264510, 264511, 264593, 60433438, 21906754, 265011, 264603, 18108351, 264288, 21906765, 21906768, 21906769, 29148629, 52644150, 264693, 33657109, 18108374, 264634, 18108385, 60432113, 22279000, 264565, 264486 265008, 265019, 264639, 22279002 |
| 2829 | 87651244 (5657, 5658) | Novel Protein sim. GBank gi 4680689 gb AAD27734.1 AF13295 - (AF13295) CGI-25 protein [Homo sapiens] | | | |
| 2830 | 88087109 (5659, 5660) | Novel Protein sim. GBank gi 2498667 sp Q61200 NPH1_MOUSE - NEUREXOPHILIN 1 | | | |
| 2831 | 87614717 (5661, 5662) | | | UNCLASSIFIED | 265017 |
| 2832 | 87631809 (5663, 5664) | | | UNCLASSIFIED | 22278997, 22278999, 52646317, 264288, 264688, 21906767, 60431528, 264638, 22279000 |
| 2833 | 87612938 (5665, 5666) | Novel Protein sim. GBank gi 5262615 emb CAB45747.1 - (AL080156) hypothetical protein [Homo sapiens] | | UNCLASSIFIED | 264555, 264556, 264558 |
| 2834 | 86974703 (5667, 5668) | Novel Protein sim. GBank gi 2224567 dbj BAA20772 - (AB002311) KIAA0313 [Homo sapiens] | | | 263972 |
| 2835 | 87775712 (5669, 5670) | Novel Protein sim. GBank gi 4589532 dbj BAA76788.1 - (AB023161) KIAA0944 protein [Homo sapiens] | | ATPase associated | 60432289, 29331828, 265008, 265010, 265017, 264448, 55811957, 265020, 18108370 |
| 2836 | 85724748 (5671, 5672) | Novel Protein sim. GBank gi 2351568 (U76618) - N-RAP [Mus musculus] | Contains protein domain (PF00412) - LIM domain containing proteins | transcript factor | 264259, 264112, 265010, 264762, 264764, 263974, 264555, 264558, 264559 |

| | | | | | |
|------|-----------------------|---|---|--------------|--|
| 2837 | 87766482 (5673, 5674) | Novel Protein sim. GBank glj5420387[embjCAB46679.1] - (AJ243459) proteophosphoglycan [Leishmania major] | | UNCLASSIFIED | 18108394, 22278997, 22278998, 29331822, 264907, 264909, 265006, 265007, 265018, 265019, 264682, 264288, 21906766, 21906767, 55811957, 35685917, 18108374, 56182323, 22278000, 22279002 |
| 2838 | 87775392 (5675, 5676) | Novel Protein sim. GBank glj973378 (U31263) - core protein [Hepatitis C virus] | | UNCLASSIFIED | 18108394, 18108397, 264259, 29331826, 265007, 265019, 264448, 18108368, 20281149, 264565, 264567 |
| 2839 | 85799317 (5677, 5678) | Novel Protein sim. GBank glj1575515 (U64899) - thrombospondin-related anonymous protein [Plasmodium gallinaceum] | | UNCLASSIFIED | 264555 |
| 2840 | 87774665 (5678, 5680) | Novel Protein sim. GBank glj2224605[dbjBAA20790] - (AB002330) KIAA0332 [Homo sapiens] | | | 264509, 264511, 265011, 264288, 264769, 265020, 264634, 264636, 264556 |
| 2841 | 86982568 (5681, 5682) | Novel Protein sim. GBank glj2224605[dbjBAA20790] - (AB002330) KIAA0332 [Homo sapiens] | Contains protein domain (PF000076) - RNA recognition motif (a.k.a. RRM, RBD, or RNP domain) | dna_rna_bind | 56182575, 35696052, 264907, 264908, 264909, 264593, 264595, 264766, 265022, 264691, 33657182, 35685763, 18108370, 35695855, 264631, 264559, 264563, 264567 |
| 2842 | 80080086 (5683, 5684) | Novel Protein sim. GBank glj578957[embjCAB51350.1] - (AL050306) dJ47587.1 (novel KIAA0615 and KIAA0323 LIKE protein) [Homo sapiens] | | UNCLASSIFIED | 264600 |
| 2843 | 91012494 (5685, 5686) | Novel Protein sim. GBank glj578957[embjCAB51350.1] - (AL050306) dJ47587.1 (novel KIAA0615 and KIAA0323 LIKE protein) [Homo sapiens] | | UNCLASSIFIED | 264906, 264907, 264908, 264909, 264910, 264784, 35695855, 83373044, 18108385 |
| 2844 | 56731154 (5687, 5688) | Novel Protein sim. GBank glj585123[spjQ08878]FBLN_MOUSE - FIBULIN-1, ISOFORM C PRECURSOR (BASEMENT-MEMBRANE PROTEIN 90) (BM-90) | Contains protein domain (PF00008) - EGF-like domain | | 264685 |
| 2845 | 94321719 (5689, 5690) | Novel Protein sim. GBank glj5420387[embjCAB46679.1] - (AJ243459) proteophosphoglycan [Leishmania major] | | homeobox | 29146498, 87168474, 264686, 35696423, 83373044, 264564 |
| 2846 | 88318613 (5691, 5692) | Novel Protein sim. GBank glj5306263[gbjAADA41995.1]AC006233 (AC006233) unknown protein [Arabidopsis thaliana] | | UNCLASSIFIED | 29331830, 264909, 265008, 265011, 87168559, 264629, 264556 |
| 2847 | 81811757 (5693, 5694) | Novel Protein sim. GBank glj3399676 (AC005390) - R31180.1 [Homo sapiens] | | UNCLASSIFIED | 264908, 264766, 264769, 264629, 264637, 264566 |
| 2848 | 87612943 (5695, 5696) | Novel Protein sim. GBank glj5262615[embjCAB45747.1] - (AL080156) hypothetical protein [Homo sapiens] | | UNCLASSIFIED | 264490, 264259, 264508, 264905, 264907, 264510, 265007, 265008, 264591, 264592, 264593, 264594, 264595, 55812038, 265010, 265011, 264604, 264763, 264764, 264765, 264766, 264686, 264628, 264629, 264555, 264636, 264556, 264557, 264638, 264558, 264559, 264563, 264566, 264567 |
| 2849 | 88084283 (5697, 5698) | Novel Protein sim. GBank glj3342218 (AC004131) - Unknown gene product [Homo sapiens] | | | 29331822, 35696052, 264509, 264906, 265007, 264594, 265018, 264288, 263972, 35696423, 18108384, 56526486, 18108390, 35696286, 265008, 265009, 265018, 264288, 35695917, 264693, 18108374, 35695855, 22279000 |
| 2850 | 87623836 (5699, 5700) | | | UNCLASSIFIED | |

| | | | | | |
|------|-----------------------|--|---|--------------|--|
| 2851 | 87820548 (5701, 5702) | Novel Protein sim. GBank gij4321619[gb AAD15788.1] - (AF051098) seven transmembrane domain orphan receptor [Mus musculus] | | UNCLASSIFIED | 264906, 264907, 56182435, 264758, 55811386, 265010, 18108351, 264448, 264369, 21906765, 264691, 264692, 264693, 35695855, 264556, 18108385, 264567, 264591 |
| 2852 | 86987023 (5703, 5704) | Novel Protein sim. GBank gij1825729 (U88308) - similar to drosophila membrane protein PATCHED SP:P18502 (PID:g129645) [Caenorhabditis elegans] | | UNCLASSIFIED | |
| 2853 | 87784630 (5705, 5706) | Novel Protein sim. GBank gij2702347 (AF027503) - putative membrane-associated guanylate kinase 1 [Mus musculus] | Contains protein domain (PF00337) - WW domain | kinase | 56182575, 55811150, 264690, 27486262, 27486285, 264632, 56182323, 56526486, 22279002 |
| 2854 | 88083557 (5707, 5708) | Novel Protein sim. GBank gij2795825 (AC004021) - kelch protein, ring canal component involved in cytoplasmic bridges; 77% Similarity to A45773 (PID:g1079096) [Homo sapiens] | Contains protein domain (PF01344) - Ketch motif | dna_ma_bind | 35696286, 29331824, 29331826, 29331828, 264908, 264768, 264693, 22279002, 264482 |
| 2855 | 94723856 (5709, 5710) | Novel Protein sim. GBank gij1504040[dj BAA13219] - (D86983) similar to D.melanogaster peroxidasin(U11052) [Homo sapiens] | Contains protein domain (PF00560) - Leucine Rich Repeat | glycoprotein | 22278994, 29331822, 29331824, 29331825, 264908, 264908, 265009, 33109954, 265018, 265019, 264448, 21906765, 265020, 264690, 27486285, 83373044, 22279000, 22279002, 264482 |
| 2856 | 88093359 (5711, 5712) | Novel Protein sim. GBank gij3264583 (AC005189) - match to ESTs H97758 (NID:g1118643) and AA085546 (NID:g1628773) [Homo sapiens] | | | 21906766, 22278997, 265022, 29331822, 29331826, 27486282, 265007, 265009, 265017, 264482, 264563, 18108351 |
| 2857 | 95348286 (5713, 5714) | Novel Protein sim. GBank gij3041855 (AC004537) - similar to tumor suppressor p33ING1; similar to AF044076 (PID:g2829208) [Homo sapiens] | Contains protein domain (PF00628) - PHD-finger | struc | 22278995, 35696286, 29331824, 29331825, 35696052, 264103, 264108, 56182435, 21906765, 21906769, 265020, 18108368, 35695763, 22279002, 264563 |
| 2858 | 87434748 (5715, 5716) | Novel Protein sim. GBank gij462585[sp P35227 ME18_HUMAN - DNA-BINDING PROTEIN MEL-18 (ZINC FINGER PROTEIN 144)] | Contains protein domain (PF00097) - Zinc finger, C3HC4 type (RING finger) | dna_ma_bind | 264369, 264887, 22278995, 22278996, 22278997, 22278999, 264259, 29331826, 29331827, 29331828, 264509, 264905, 264906, 29331830, 264908, 52644045, 264909, 264511, 264512, 265007, 265008, 264910, 265009, 264593, 60433356, 264595, 264758, 21906754, 265010, 265011, 264604, 265018, 264760, 18108351, 264763, 264682, 264764, 264765, 264288, 264369, 264685, 264766, 264768, 18108357, 264769, 21906766, 21906767, 265021, 264534, 60170615, 264691, 264692, 18108370, 264629, 18108374, 264631, 264636, 263981, 18108381, 264558, 18108385, 22279002, 264564, 264566, 264486, 264567 |
| 2859 | 80937675 (5717, 5718) | Novel Protein sim. GBank gij4325320[gb AAD17331.1] - (AF124427) claudin-15 [Mus musculus] | | UNCLASSIFIED | 60424179, 65274572, 29331828, 264905, 264511, 264758, 265011, 21906767, 21906769, 55811957, 265021, 56182323 |

| | | | | | |
|------|-----------------------|--|---|--------------|--|
| 2860 | 87532589 (5719, 5720) | Novel Protein sim. GBank gi 4469186 emb CAB38414.1 - (AL031588) dJ1163J1.2.1 (novel protein similar to C. elegans B0035.16 and bacterial RNA (5'-Methylaminomethyl-2-thiouridylyl)-Methyltransferases) (isoform 1) [Homo sapiens] | | UNCLASSIFIED | 284259, 29331822, 29331824, 29331825, 29331826, 29331827, 29331828, 264510, 264511, 33109954, 18108351, 264683, 264765, 264369, 264686, 21906765, 264691, 264692, 264693, 18108388, 22279002, 264482 |
| 2861 | 86698507 (5721, 5722) | Novel Protein sim. GBank gi 3941730 (AF108083) - BS4 [Homo sapiens] | | | 264369, 264692 |
| 2862 | 87569585 (5723, 5724) | Novel Protein sim. GBank gi 4505013 ref NP_002310.1 pLRN - leucine-rich neuronal protein | | UNCLASSIFIED | 264691, 264638 |
| 2863 | 91220421 (5725, 5726) | Novel Protein sim. GBank gi 3249088 (AC004473) - Contains similarity to goliath protein gb M97204 from D. melanogaster [Arabidopsis thaliana] | Contains protein domain (PF000036) - struct Zinc finger, C2H2 type | | 55994075, 35686286, 22278998, 29331822, 29331824, 35686052, 29331828, 264106, 264511, 55812038, 33657084, 55811386, 265018, 265019, 21906765, 21906766, 21906769, 35695917, 265020, 265022, 33657023, 33657109, 33657349, 264629, 18108376, 60431850, 56182323, 18108385, 18108387, 87168518, 22279002 |
| 2864 | 87420030 (5727, 5728) | Novel Protein sim. GBank gi 1079451 pir JA55463 - tropomodulin, skeletal muscle - chicken | | struct | 264259, 284910 |
| 2865 | 95312191 (5729, 5730) | Novel Protein sim. GBank gi 438840 (L19048) - MSA-2 [Plasmodium falciparum] | | | 22278995, 21906764, 264482 |
| 2866 | 95105480 (5731, 5732) | Novel Protein sim. GBank gi 585703 sp Q07066 PMP2_RAT - 22 KD PEROXISOMAL MEMBRANE PROTEIN | | glycoprotein | 65274572, 22278996, 22278998, 22278999, 264259, 29331824, 29331827, 29331828, 60433438, 21906754, 265018, 264448, 264764, 52644150, 83373044 |
| 2867 | 86908001 (5733, 5734) | Novel Protein sim. GBank gi 4580997 gb AAD24571.1 AF12108 - (AF121081) cAMP inducible 2 protein [Mus musculus] | | UNCLASSIFIED | 264488, 264768, 21906768, 22278998, 265022, 264259, 264508, 264905, 264907, 264511, 264910, 264635, 264636, 264637, 265011, 265017, 265018, 265019, 264583, 264088, 264566, 264764, 264369, 264567, 264486, 264288, 264766 |

| | | | | |
|------|-----------------------|---|--------------|--|
| 2868 | 95303283 (5735, 5736) | Novel Protein sim. GBank gij1202868[emb]CAA63923] - (X94232) t-Cell activation protein [Homo sapiens] | | 18108392, 264488, 22278994, 22278997, 22278998, 22278999, 60432049, 264259, 29331822, 29331824, 29331825, 60432289, 29331826, 29331827, 35696052, 29331828, 264508, 52844045, 264828, 265006, 265007, 265008, 265009, 264591, 60432229, 264593, 60433358, 264595, 21906754, 285017, 265019, 264682, 264369, 21906765, 21906766, 21906767, 21906768, 265021, 265022, 264691, 33657182, 18108388, 27486261, 27486262, 27486264, 27486265, 18108370, 18108374, 35696423, 35695855, 264632, 56182323, 87168518, 264404, 22278000, 22279002, 264482, 264563, 264564, 264567, 264487 |
| 2869 | 88094412 (5737, 5738) | | UNCLASSIFIED | 264369 |
| 2870 | 84404574 (5739, 5740) | | UNCLASSIFIED | 264905, 264908, 264764, 21906769, 264634 |
| 2871 | 88318621 (5741, 5742) | Novel Protein sim. GBank gij5306263[gb]AAD41995.1[AC006233] unknown protein [Arabidopsis thaliana] | UNCLASSIFIED | 264259, 29331822, 60432289, 29331827, 264907, 265008, 265017, 265018, 264682, 264764, 18108354, 265021, 27486265, 264629, 18108387, 264567 |
| 2872 | 95312187 (5743, 5744) | Novel Protein sim. GBank gij112205[pir]B39066 - proline-rich protein 15 - rat | kinase | 263981 |
| 2873 | 88094252 (5745, 5746) | | UNCLASSIFIED | 264488, 18108374, 264768, 264687, 264688, 264689, 35696423, 35696288, 35695917, 264510, 264511, 265007, 264512, 265008, 264910, 264534, 264634, 264635, 264555, 264592, 264259, 264558, 60433438, 60432289, 35696052, 265011, 264600, 264601, 60432113, 264508, 264563, 264482, 264509, 264905, 264908, 264564, 18108351, 264763, 18108370, 264907, 264566, 264908, 264764, 264288, 264567, 264909, 264488, 264766, 18108391 |

| | | | | |
|------|-----------------------|---|---|--|
| 2874 | 94313548 (5747, 5748) | Novel Protein sim. GBank glj3212854 (AC004005) - unknown protein [Arabidopsis thaliana] | UNCLASSIFIED | 52644507, 52645156, 52646355, 56182575, 22278994, 22278995, 56994075, 35696286, 60432049, 264259, 52645080, 29331822, 56182181, 29331824, 80424269, 29331825, 66714117, 29331826, 29331827, 35696052, 29331828, 33656970, 66712502, 52644045, 285007, 265009, 60433356, 264758, 55812038, 18108348, 52646317, 33109954, 33657084, 265017, 264604, 285018, 285019, 264682, 264369, 264288, 264688, 52644229, 21906766, 21908767, 21906768, 55811957, 265020, 265021, 265022, 52644150, 33657023, 52845129, 18108374, 18108376, 35696423, 56182323, 18108387, 87168518, 60432113, 22279000, 22279002, 264563, 264565 |
| 2875 | 88083726 (5749, 5750) | Novel Protein sim. GBank glj2781386 (AC004010) - similar to Leucine-rich transmembrane proteins; 44% similarity to U42767 (PID:g1736918) [Homo sapiens] | glycoprotein | 22278996, 22278997, 22278999, 29331826, 29331828, 29146499, 66712502, 265008, 265017, 18108351, 264683, 264689, 21908767, 18108376, 18108377, 55811576, 60170394, 22279000, 264487 |
| 2876 | 88080854 (5751, 5752) | Novel Protein sim. GBank glj2979530 (AC004449) - R33683_2 [Homo sapiens] | Contains protein domain (PF00167) - fgl | |
| 2877 | 94747029 (5753, 5754) | Novel Protein sim. GBank glj4704208 [emb] (CAB41646.1) - (AL035419) dJ1100H13.1 (putative novel protein) [Homo sapiens] | | 52646365, 65274572, 56182575, 22278997, 22278998, 22278999, 60432049, 52645080, 60424269, 60432289, 29331827, 35696052, 29331828, 66712502, 52644045, 56182435, 60433356, 33657402, 33657084, 265019, 55811150, 264448, 264388, 21908766, 21906768, 21906769, 265020, 33657023, 33657109, 33657182, 27486262, 264629, 60431528, 55811576, 52644332, 56182323, 264558, 83373044, 18108385, 56526486, 60432113, 22279000, 22279002, 264563 |
| 2878 | 88095309 (5755, 5756) | Novel Protein sim. GBank glj387677 [emb] (CAB03067) - (Z81077) predicted using GeneFinder; Similarity to Yeast protein 8248 (TR:G587531) [Caenorhabditis elegans] | UNCLASSIFIED | 264488, 264258, 29331822, 29331826, 264805, 264509, 264907, 264909, 264510, 265005, 264511, 264512, 33657402, 264758, 21908754, 18108351, 264681, 264682, 264288, 264684, 264685, 264768, 264689, 21908769, 264690, 33657023, 264693, 18108364, 33657109, 264629, 18108374, 264630, 264632, 264556, 264637, 264639, 264558, 18108385, 18108387, 264563, 264564, 264565, 264566, 264486, 264567 |

| | | | | | |
|------|-----------------------|--|--|------------------|--|
| 2879 | 87859122 (5757, 5758) | Novel Protein sim. GBank gi 4895145 gb AAD32752.1 - (AF127374) unknown [Streptomyces lavendulae] | Contains protein domain (PF00315) - Uracil-DNA glycosylase | UNCLASSIFIED | 18108359, 264259, 264905, 18108370, 264629, 264908, 264909, 18108374, 18108377, 265008, 264910, 264637, 60170394, 264559, 265017, 264564, 264585, 264587, 264684, 264369 |
| 2880 | 94851439 (5759, 5760) | Novel Protein sim. GBank gi 4680703 gb AAD27741.1 AF13296 - (AF132966) CGI-32 protein [Homo sapiens] | | | 264488, 52646365, 52646842, 22278994, 35696286, 22278998, 22278999, 264259, 29331822, 29331824, 29331825, 29331826, 29331827, 35696052, 29331828, 264107, 264508, 264509, 264805, 264906, 264907, 264908, 264909, 52644045, 264510, 265006, 264511, 265007, 264512, 265009, 264910, 264594, 21908754, 52646317, 52644298, 87168559, 264600, 264604, 264605, 264760, 264764, 264288, 264766, 264768, 264687, 264769, 21908766, 21908769, 35695917, 265021, 264690, 264692, 33657023, 52645129, 33657109, 33657182, 27486262, 33657349, 264629, 18108374, 35695855, 264634, 264635, 264636, 264637, 264638, 264557, 52644332, 264558, 264559, 83373044, 264404, 22279000, 264563, 264483, 264567, 264486 |
| 2881 | 87650539 (5761, 5762) | Novel Protein sim. GBank gi 733571 (U23452) - No definition line found [Caenorhabditis elegans] | | UNCLASSIFIED | 22278998, 29331822, 52644045, 21906765, 264639, 60432113 |
| 2882 | 87714387 (5763, 5764) | Novel Protein sim. GBank gi 1118112 (U41559) - No definition line found [Caenorhabditis elegans] | | | 264488, 22278998, 22278999, 29331822, 29331826, 264908, 60170831, 60433356, 55812038, 264681, 264682, 264686, 264687, 264688, 21908768, 21908769, 264693, 263987, 18108374, 55811578, 56182323, 22279002, 264566 |
| 2883 | 95362875 (5765, 5766) | Novel Protein sim. GBank gi 4686008 gb AAD31087.1 AF10693 - (AF106934) vitamin D receptor-interacting protein [Homo sapiens] | Contains protein domain (PF00400) - WD domain, G-beta repeat | UNCLASSIFIED | 264112, 264682 |
| 2884 | 87784643 (5767, 5768) | Novel Protein sim. GBank gi 2224697 dbj BAA20832 - (AB002376) KIAA0378 [Homo sapiens] | | UNCLASSIFIED | 265018, 264634 |
| 2885 | 83006308 (5769, 5770) | Novel Protein sim. GBank gi 1255889 (U53344) - T07H6.5 gene product [Caenorhabditis elegans] | | UNCLASSIFIED | 264686, 264693 |
| 2886 | 91237823 (5771, 5772) | Novel Protein sim. GBank gi 3882323 dbj BAA34521.1 - (AB018344) KIAA0801 protein [Homo sapiens] | Contains protein domain (PF00084) - Sushi domain (SCR repeat) | complementrecept | 60432049, 264259, 29331828, 264908, 264511, 264595, 60433438, 264596, 265017, 264605, 263989, 263972, 264555, 83373044, 87168518, 264566 |
| 2887 | 91227860 (5773, 5774) | Novel Protein sim. GBank gi 3882323 dbj BAA34521.1 - (AB018344) KIAA0801 protein [Homo sapiens] | | UNCLASSIFIED | 18108351, 264686, 264629, 264631, 264639, 83373044, 264482 |
| 2888 | 95105816 (5775, 5776) | Novel Protein sim. GBank gi 4508015 ref NP_003447.1 pZNF2 - zinc finger protein 205 [Homo sapiens] | Contains protein domain (PF00096) - Zinc finger, C2H2 type | transcriptfactor | 264488, 264259, 29331828, 264508, 264906, 264593, 264758, 264766, 264769, 18108374, 83373044, 264486 |

| | | | | | |
|------|-----------------------|---|--|---------------------|--|
| 2889 | 87606562 (5777, 5778) | | | UNCLASSIFIED | 56994075, 29331824, 265009, 264760, 18108354, 264288 |
| 2890 | 78703853 (5779, 5780) | Novel Protein sim. GBank gij854085[emb CAA58337] - [X83413] U89 [Human herpesvirus 6] | | UNCLASSIFIED | 264591, 264766 |
| 2891 | 88094428 (5781, 5782) | Novel Protein sim. GBank gij3877750[emb CAB01508] - (Z78084) predicted using GeneFinder; similar to collagen; cDNA EST EMBL:D65865 comes from this gene; cDNA EST EMBL:D69451 comes from this gene; cDNA EST EMBL:D66026 comes from this gene; cDNA EST EMBL:D69658 comes from this gene... | | UNCLASSIFIED | 264591, 264595, 264369, 264685, 264693, 264628, 264563, 264566 |
| 2892 | 95418745 (5783, 5784) | Novel Protein sim. GBank gij4929759[gb AAD34140.1]AF15190 - (AF151903) CGI-145 protein [Homo sapiens] | | UNCLASSIFIED | 65274572, 35696286, 29331828, 264110, 265009, 60433438, 265018, 265019, 18108351, 264288, 21906765, 21906766, 21906769, 55811957, 264690, 65274620, 263967, 35695855 |
| 2893 | 87798014 (5785, 5786) | | | UNCLASSIFIED | 264259, 265010, 264682, 18108370, 264555, 264556 |
| 2894 | 87755985 (5787, 5788) | Novel Protein sim. GBank gij5668015[gb AAD46135.1] - (AF080171) zinc finger protein ZNF232 [Homo sapiens] | Contains protein domain (PF000096) - Zinc finger, C2H2 type | - transcript factor | 264259, 265006, 60433438, 52644286, 265011, 264369, 35695917, 18108381, 18108382, 18108388 |
| 2895 | 86938778 (5789, 5790) | Novel Protein sim. GBank gij3924708[emb CAA84646] - (Z35597) Weak similarity with sea squirt nidogen precursor protein (blastp score 71); cDNA EST EMBL:T02069 comes from this gene; cDNA EST EMBL:D76135 comes from this gene; cDNA EST EMBL:D73147 comes from this gene; cDNA EST EMB... | Contains protein domain (PF01437) - Plexin repeat | | 29331824, 265007, 264762, 264636, 264563 |
| 2896 | 87752122 (5791, 5792) | Novel Protein sim. GBank gij4885549[ref NP_005456.1]pKRG - protein kinase B gamma | Contains protein domain (PF00089) - Eukaryotic protein kinase domain | kinase | 18108392, 18108394, 18108398, 22278998, 264259, 29331822, 29331824, 29146499, 264906, 264908, 265007, 265009, 265018, 265019, 264369, 264685, 264689, 21906766, 265021, 264693, 33657182, 264639, 18108384, 18108388, 264567 |
| 2897 | 95413057 (5793, 5794) | Novel Protein sim. GBank gij4502877[ref NP_001296.1]pCLDN - Clostridium perfringens enterotoxin receptor 1 | | UNCLASSIFIED | 60424179, 56182575, 22278996, 35696286, 22278997, 22278999, 60432049, 264259, 29331824, 29331825, 60424269, 60432289, 29331826, 29331828, 35696052, 264908, 56182435, 265009, 264910, 60170831, 60431735, 60433356, 60433438, 65274444, 55811388, 265018, 18108351, 264448, 264288, 264687, 21906765, 21906766, 21906767, 21906769, 55811857, 35695917, 264534, 33657023, 33657109, 35695763, 264628, 264629, 60431528, 18108374, 55810764, 55811576, 35696423, 35695855, 264555, 56182323, 18108385, 264404, 22279000, 22279002, 264566 |
| 2898 | 87750340 (5795, 5796) | | | UNCLASSIFIED | 22278995, 264604, 18108385, 264566 |

| | | | | | | |
|------|-----------------------|---|--|---|--------------|--|
| 2899 | 80357670 (5787, 5788) | | | | UNCLASSIFIED | 264764, 21906764, 264692 |
| 2900 | 94233538 (5789, 5800) | Novel Protein sim. GBank gi 4581470 emb CAB40137.1 - (Y18483) SLC7A8 protein [Homo sapiens] | | | glycoprotein | 85274572, 56182575, 35696286, 60432049, 264259, 29331824, 66714117, 29331826, 35696052, 29331828, 66712502, 56182435, 265006, 265007, 265008, 265009, 60433356, 264758, 265018, 264764, 264765, 264288, 264768, 21906764, 21906768, 21906769, 265020, 264692, 264693, 32833986, 264631, 83373044, 60432113 |
| 2901 | 87444731 (5801, 5802) | Novel Protein sim. GBank gi 4759272 ref NP_004614.1 pTTC4 - tetrairicopeptide repeat domain 4 | | | phosphatase | 22278995, 22278997, 22278999, 60432049, 29331822, 29331824, 29331825, 29331827, 35696052, 33656970, 264910, 265009, 21908754, 33657084, 87168474, 265010, 265018, 21906764, 21906765, 21906766, 21906767, 21906769, 33657023, 264693, 33657109, 33657349, 35696423, 35695855, 283981, 56182323, 22278902 |
| 2902 | 85745271 (5803, 5804) | Novel Protein sim. GBank gi 2414615 emb CAB16364 - (Z89259) hypothetical protein [Schizosaccharomyces pombe] | | | | 264683, 264691 |
| 2903 | 87608733 (5805, 5806) | Novel Protein sim. GBank gi 1079318 pir J552241 - XLCL2 protein - African clawed frog | | | | 264887, 22278994, 264259, 29331826, 29331828, 264905, 52644045, 56182435, 264511, 265017, 265018, 18108351, 264448, 264683, 264769, 264689, 35695917, 52644150, 87168518, 60432113, 22278902 |
| 2904 | 86458072 (5807, 5808) | Novel Protein sim. GBank gi 5639823 gb AAD45885.1 AF14367 - (AF143676) multispinning nuclear envelope membrane protein nurim [Homo sapiens] | | | UNCLASSIFIED | 52646365, 22278999, 264259, 35696052, 265011, 265017, 264683, 264769, 35695917, 265020, 263967, 18108374, 35695855, 284637, 264952, 18108385, 18108387 |
| 2905 | 84449826 (5809, 5810) | Novel Protein sim. GBank gi 728837 sp P39194 ALU7_HUMAN - !!!!! ALU SUBFAMILY SQ WARNING ENTRY !!!! | | | oncogene | 265009, 264681, 264682 |
| 2906 | 95341051 (5811, 5812) | Novel Protein sim. GBank gi 4689256 gb AAD27831.1 AF12185 - (AF121858) sorting nexin 8 [Homo sapiens] | | Contains protein domain (PF00787) - PX domain | UNCLASSIFIED | 22278996, 35696286, 22278998, 264259, 60432289, 29331828, 29331830, 66712502, 265009, 60170831, 33109954, 264448, 264683, 264288, 264689, 21906766, 21906767, 21906768, 55811957, 35695917, 265022, 52644150, 264691, 33657023, 264692, 264693, 35695855, 60432113, 264566 |

| | | | | |
|------|-----------------------|---|---|--|
| 2807 | 91211383 (5813, 5814) | Novel Protein sim. GBank gij1707078 (U80451) - contains strong similarity to a DNAJ-like domain (PS:PS00636) [Caenorhabditis elegans] | Contains protein domain (PF00226) - eph DnaJ domain | 52644507, 56182575, 56181686, 22278995, 56994075, 35696286, 60432049, 56182181, 35696052, 60431735, 264595, 55812038, 21906754, 55811386, 265019, 264682, 264369, 56181562, 21906766, 55811957, 35695917, 265020, 265021, 33657023, 33657109, 60431528, 55811576, 35696423, 35695855, 264638, 22279000 |
| 2808 | 80414246 (5815, 5816) | Novel Protein sim. GBank gij2673917 (AC002561) - putative ATP-dependent RNA helicase [Arabidopsis thaliana] | helicase | 265009, 33109954, 18108351, 264786, 265021, 264691, 264692, 18108374, 264556, 264638, 264557, 264558 |
| 2809 | 87420225 (5817, 5818) | | eph | 264259, 87168474, 265018, 18108365, 264628 |
| 2910 | 86601075 (5819, 5820) | Novel Protein sim. GBank gij4539335 [emb]CAB37483.1] - (AL035539) putative protein [Arabidopsis thaliana] | | 22278995, 264509, 264512, 265007, 33657402, 265017, 264369, 265022, 18108365, 264628 |
| 2911 | 94216615 (5821, 5822) | Novel Protein sim. GBank gij4469187 [emb]CAB38415.1] - (AL031588) dJ1163J1.3 (novel protein similar to mouse B99) [Homo sapiens] | glucoamylase | 52646365, 18108397, 22278995, 22278997, 22278998, 22278999, 29331824, 29331825, 52644045, 265006, 265018, 264448, 21906765, 21906766, 21906767, 21906768, 21906769, 55811957, 265021, 18108370, 18108372, 18108374, 22279000 |
| 2912 | 87731803 (5823, 5824) | Novel Protein sim. GBank gij4929637 [gb]AAD34079.1 [AF151842] CGI-84 protein [Homo sapiens] | Contains protein domain (PF00904) - Involucrin repeat | 52645136, 264092, 60432049, 264259, 52645080, 29331824, 29331825, 66712502, 33109954, 264760, 264683, 264288, 264686, 265021, 264693, 18108368, 263978, 264404 |
| 2913 | 87713823 (5825, 5826) | Novel Protein sim. GBank gij854065 [emb]CAA58337] - (X83413) U88 [Human herpesvirus 6] | UNCLASSIFIED Zinc finger, C2H2 type | 52644507, 52645156, 52646842, 56182575, 35696286, 22278997, 264259, 52645080, 29331827, 35696052, 29331828, 264828, 52644045, 56182435, 55812038, 52646317, 21906754, 52644296, 87168474, 265017, 265018, 265019, 18108351, 264682, 264686, 264689, 21906765, 21906766, 21906767, 21906768, 35695917, 265020, 52644150, 27486261, 27486262, 27486265, 35695763, 55811576, 35695855, 52644332, 22279000, 22279002, 264563 |
| 2914 | 87797300 (5827, 5828) | | | 264557 |

| | | | | |
|------|-----------------------|--|---|--|
| 2915 | 86081072 (5829, 5830) | Novel Protein sim. GBank gjl5174485[refNP_008030.1]pKIAA - endocytic receptor (macrophage mannose receptor family) | Contains protein domain (PF00059) - eph Lectin C-type domain | 264569, 264488, 264687, 264768, 21906768, 52646842, 21906787, 21906788, 56182575, 29148629, 35695917, 22278996, 22278997, 22278998, 265021, 22278999, 52644150, 264691, 264259, 60432049, 264692, 52645129, 33657109, 33657182, 29331827, 27486261, 35696052, 29331828, 27486262, 27486264, 27486265, 33657349, 29146498, 29146499, 264906, 264907, 18108370, 264908, 18108372, 52644045, 18108374, 56192435, 35695855, 264112, 264510, 265008, 60432229, 264593, 60433356, 56192323, 18108382, 55812038, 18108385, 33109954, 21906754, 33657084, 87168518, 265010, 265011, 60432113, 265017, 265018, 22278000, 265019, 55811150, 264681, 18108351, 264763, 264448, 264683, 264566, 18108354, 264389, 264288, 264766 |
| 2916 | 95337790 (5831, 5832) | Novel Protein sim. GBank gjl5104851[dbjBAA80165.1] - (AP000061) 305aa long hypothetical dTDP-4- dehydrohamnose reductase [Aeropyrum pernix] | dehydrogenase | 52645156, 65274572, 22278994, 22278995, 35696286, 22278996, 22278997, 22278998, 22278999, 264259, 29331822, 29331824, 29331825, 66714117, 60432288, 29331826, 29331827, 29331828, 33656970, 264509, 264906, 29331830, 52644045, 264909, 56182435, 60170831, 264592, 264593, 33657402, 60433356, 52646317, 21906754, 33109954, 33657084, 52644296, 85658542, 265011, 265017, 265018, 265019, 18108351, 264448, 264288, 52644229, 21906765, 21906766, 21906767, 21906768, 21906769, 55811957, 35695917, 265021, 265022, 52644150, 33657023, 33657109, 33657182, 27486261, 27486262, 27486264, 35695763, 18108376, 55811576, 35696423, 85274791, 35695855, 52644332, 264557, 264638, 56182323, 18108387, 87168518, 22278002, 264482 |
| 2917 | 87454546 (5833, 5834) | Novel Protein sim. GBank gjl3169065[embjCAA19260.1] - (AL023704) putative translocation elongation factor-Tu fa mily [Schizosaccharomyces pombe] | UNCLASSIFIED | 60433438, 264602, 264682, 87168518, 60432113 |
| 2918 | 85690529 (5835, 5836) | Novel Protein sim. GBank gjl539218[pirljS38038 - hypothetical protein YKL201c - yeast (Saccharomyces cerevisiae)] | | 264638 |
| 2919 | 87641497 (5837, 5838) | Novel Protein sim. GBank gjl2564855 (AF030001) - unknown [Mus musculus] | | 66714117, 66712502, 263981 |

| | | | | | |
|------|-----------------------|--|---|---------------|--|
| 2920 | 87769523 (5839, 5840) | | | | 35686286, 22278987, 264259, 52645080, 29331824, 29331826, 29331827, 264828, 264909, 56182435, 264511, 264758, 33109954, 21906754, 52644296, 265010, 265011, 264601, 265017, 265019, 264681, 264687, 21906767, 265021, 52644150, 264690, 264691, 264692, 264693, 33657109, 33657182, 27486262, 27486284, 27488265, 35696423, 35695855, 264632, 264636, 264637, 264638, 56182323, 60170394, 18108385, 87168518, 60432113 |
| 2921 | 91639882 (5841, 5842) | Novel Protein sim. GBank glj4580013jgb/AAD24202.1 U83194_ - (U83194) TRAF4-associated factor 2 [Homo sapiens] | Contains protein domain (PF00787) - PX domain | | 35686286, 22278987, 264091, 264092, 264094, 264259, 29331822, 29331824, 29331826, 29331827, 35696052, 29146498, 264104, 264105, 264107, 264509, 264110, 264112, 264512, 60433356, 21906754, 87168474, 265017, 18108351, 264288, 21906765, 21906766, 21906767, 21906769, 35695917, 265021, 263974, 18108374, 263976, 263977, 18108376, 264555, 263981, 56526486, 87168518, 22279000, 22279002 |
| 2922 | 87749762 (5843, 5844) | Novel Protein sim. GBank glj4589514 dbj BAA76779.1 - (AB023152) KIAA0935 protein [Homo sapiens] | Contains protein domain (PF01074) - Glycosyl hydrolases family 38 | kinase | 264908, 264909, 264511, 265006, 265008, 264593, 33657402, 60174639, 18108351, 264763, 21906765, 29148627, 35695917, 264692, 264629, 263978, 55811576, 35695855, 264555, 264558, 56182323, 60170394, 22279000, 264486 |
| 2923 | 95337789 (5845, 5846) | Novel Protein sim. GBank glj4835268 emb CAB42898.2 - (Z83844) dJ37E16.4 (similar to mouse p116Rip protein) [Homo sapiens] | Contains protein domain (PF00169) - PH domain | struct | 264488, 18108397, 22278985, 22278986, 22278987, 22278998, 22278999, 29331825, 29331826, 29331827, 29331830, 264511, 265009, 33657402, 265011, 265017, 265018, 264683, 18108354, 21906765, 21906767, 21906768, 21906769, 52644150, 264691, 264692, 33657109, 263974, 18108376, 264631, 264636, 18108385, 18108387, 22279000, 264563, 264566 |
| 2924 | 87781967 (5847, 5848) | Novel Protein sim. GBank glj2133095 pir S72254 - ribosomal protein L36, mitochondrial - yeast (Saccharomyces cerevisiae) | Contains protein domain (PF00444) - Ribosomal protein L36 | ribosomalprot | 265017, 264628, 20281152, 264556 |
| 2925 | 95090120 (5849, 5850) | Novel Protein sim. GBank glj2388986 emb CAB11718 - (Z89880) actin associated protein [Schizosaccharomyces pombe] | UNCLASSIFIED | | 56182575, 35686286, 264259, 60432289, 29331827, 264508, 52644045, 264910, 264591, 60432229, 55812038, 21906754, 264681, 264448, 264683, 264288, 264685, 52644229, 264689, 21906765, 21906768, 21906769, 265021, 265022, 60170615, 264692, 33657023, 264693, 33657109, 35686423, 65274791, 56182323 |

| | | | | | |
|------|-----------------------|--|---|--------------|--|
| 2926 | 85343003 (5851, 5852) | | | | 29331828, 265011, 264768, 264689 |
| 2927 | 80408018 (5853, 5854) | Novel Protein sim. GBank gi 283032 pir S22456 - hydroxyproline-rich glycoprotein - perennial teosinte | | | 264764, 264288, 264630, 264637 |
| 2928 | 20452179 (5855, 5856) | | | UNCLASSIFIED | 264559 |
| 2929 | 81622020 (5857, 5858) | Novel Protein sim. GBank gi 3413320 emb CAA069151 - (AJ006215) CMP-N-acetylneuraminic acid synthetase [Mus musculus] | | UNCLASSIFIED | 264589, 264489, 22278994, 35696286, 22278996, 22278998, 22278999, 264084, 264259, 52645080, 29331822, 29331824, 66714117, 29331825, 29331826, 29331827, 35696052, 33656970, 264109, 29331830, 52644045, 285009, 33109954, 52644296, 87168559, 264760, 264762, 264448, 264764, 264288, 264766, 264768, 21906765, 21906766, 21906768, 21906769, 35695917, 264691, 33657023, 264693, 33657109, 18108374, 263976, 35698423, 35695855, 263981, 22279000, 22279002, 264567, 264486 |
| 2930 | 95302755 (5859, 5860) | | | UNCLASSIFIED | 56182575, 56181686, 35696286, 22278996, 22278998, 22278999, 264259, 29331825, 60432289, 29331828, 264905, 52644045, 56182435, 285009, 60170831, 264592, 60432229, 60433356, 87168474, 285010, 265011, 265017, 265018, 265019, 264762, 264448, 264683, 264288, 264768, 21906765, 21906769, 35695917, 60170615, 33657023, 33657109, 264628, 18108370, 18108372, 35696423, 35695855, 264556, 56182323, 60432113, 264567 |
| 2931 | 94312693 (5861, 5862) | Novel Protein sim. GBank gi 3788433 (AF098505) - similar to Arabidopsis thaliana male sterility protein 2 (SW:Q08891) [Caenorhabditis elegans] | Contains protein domain (PF00471) - Ribosomal protein L33 | UNCLASSIFIED | 52645156, 22278997, 22278998, 29331822, 52645080, 29331824, 60432289, 33656970, 60433356, 60433438, 33109954, 21906765, 21906766, 21906767, 21906768, 265020, 52644150, 33657023, 33657109, 33657182, 27488265, 35696423, 35695855, 264555, 87168518, 60432113, 264566 |
| 2932 | 79632623 (5863, 5864) | | | | 264906, 264907 |
| 2933 | 91720776 (5865, 5866) | Novel Protein sim. GBank gi 3378056 (AF017777) - helicase [Drosophila melanogaster] | helicase | | 264488, 18108392, 56182575, 22278999, 264091, 264259, 29331825, 60432289, 29331827, 264508, 52644045, 56182435, 265007, 265009, 264592, 60433356, 60433438, 21906754, 265017, 264682, 264288, 52644229, 21906765, 21906766, 21906769, 265022, 52644150, 33657023, 33657109, 27488265, 264635, 264636, 60170394, 56182323, 18108385, 60432113, 264565, 264566, 264567 |

| | | | | | |
|------|-----------------------|--|--|-------------------|---|
| 2934 | 86576025 (5867, 5868) | | | | 22278997, 22278999, 29331824, 33657402, 264691, 27486282, 264628, 87168518, 22279000 |
| 2935 | 86410579 (5869, 5870) | | | UNCLASSIFIED | 56182575, 22278995, 60433356, 33657402, 264758, 33109954, 21908754, 265018, 265019, 264448, 264769, 21906764, 21906765, 265021, 284692, 33657023, 33657109, 33657349, 55810764, 22279000 |
| 2936 | 87605863 (5871, 5872) | Novel Protein sim. GBank gi 4153862 (AC005065) - determined by GENSCAN prediction and spliced EST; match to EST R84329 (NID:942735) [Homo sapiens] | Contains protein domain (PF00856) - SET domain | nuclease | 22278997, 29331827, 29331828, 265009, 265017, 264605, 265020, 55811576, 18108387, 60432113, 264563 |
| 2937 | 94853086 (5873, 5874) | Novel Protein sim. GBank gi 5174409[ref NP_006101.1 pCD2B - CD2 antigen (cytoplasmic tail)-binding protein 2 | | UNCLASSIFIED | 56994075, 22278999, 264259, 604332049, 29331822, 56182181, 29331827, 29331828, 264906, 264908, 264909, 56182435, 265006, 264512, 264910, 60170831, 60433356, 265011, 265018, 18108351, 264448, 264288, 264766, 52644229, 21906765, 29148784, 65274791, 264556, 56182323, 60170394, 264558, 60432113, 264565, 264486, 264587 |
| 2938 | 95419773 (5875, 5876) | Novel Protein sim. GBank gi 3319990[emb CAA76720] - (Y17287) ubiquitin-conjugating enzyme [Mus musculus] | Contains protein domain (PF00179) - Ubiquitin-conjugating enzyme | ubiquitin | 264488, 56182575, 22278996, 35696286, 22278997, 22278998, 22278999, 264490, 264259, 29331822, 29331824, 6671417, 29331827, 35696052, 264107, 264905, 66712502, 52644045, 56182435, 264511, 265008, 265009, 60432229, 33657402, 60433438, 55812038, 21906754, 85658542, 265010, 265011, 87168559, 265017, 265018, 265019, 264681, 264288, 264689, 21906765, 21906767, 21906768, 55811957, 35695917, 265020, 60170615, 264690, 264691, 264692, 33657023, 264693, 65274620, 33657109, 18108370, 18108374, 263976, 35696423, 35695855, 264555, 264556, 18108381, 56182323, 60170394, 83373044, 18108385, 56526486, 60432113, 22279002 |
| 2939 | 87786622 (5877, 5878) | Novel Protein sim. GBank gi 3979900[emb CAA99909] - (Z75547) similar to WD domain, G-beta repeat; cDNA EST yk371b7.5 comes from this gene; cDNA EST yk312h1.5 comes from this gene; cDNA EST yk465d5.5 comes from this gene; cDNA EST yk472c4.5 comes from this gene; cDNA EST yk292f8.... | Contains protein domain (PF00400) - WD domain, G-beta repeat | ATPase-associated | 264907, 265018, 264681, 264685, 264686 |

| | | | | | | |
|------|-----------------------|--|---|--|-------------------|--|
| 2940 | 95011103 (5879, 5880) | | | | UNCLASSIFIED | 22278996, 29331822, 29331824, 66714117, 29331826, 29331828, 264905, 264908, 66712502, 29331830, 265011, 265017, 264764, 264369, 21906766, 21906767, 33657023, 33657109, 32833986, 18108374, 18108377, 264634, 83373044, 18108385, 18108387, 264566 |
| 2941 | 21423370 (5881, 5882) | Novel Protein sim. GBank gi 3413872 dbj BAA32300 - (AB007924) KIAA0455 protein [Homo sapiens] | | | UNCLASSIFIED | 264557 |
| 2942 | 87430203 (5883, 5884) | Novel Protein sim. GBank gi 1172845 sp P46629 RB25_RABIT - RAS-RELATED PROTEIN RAB-25 | | | glycoprotein | 264910, 265010, 264768 |
| 2943 | 95314504 (5885, 5886) | Novel Protein sim. GBank gi 4928653 gb AAD34087.1 AF15185 - (AF151850) CGI-82 protein [Homo sapiens] | | | collagen | 60432049, 264259, 60432289, 29331827, 29146498, 265008, 264593, 60433356, 60433438, 265010, 265011, 265017, 265018, 264683, 264766, 18108381, 65274727, 60432113, 264567 |
| 2944 | 95081063 (5887, 5888) | Novel Protein sim. GBank gi 4678282 emb CAB41190.1 - (AL049860) 1-acylcerol-3-phosphate acyltransferase-like protein [Arabidopsis thaliana] | Contains protein domain (PF00415) - Regulator of chromosome condensation (RCC1) | | ATPase_associated | 56994075, 22278998, 60432049, 264259, 29331822, 29331824, 60424269, 60432289, 29331826, 29331828, 264905, 264907, 52644045, 264909, 264511, 265008, 265009, 264594, 21906754, 87168559, 264603, 265017, 265018, 18108351, 264682, 264766, 264687, 264689, 21908765, 21906766, 21906767, 21906768, 21906769, 265021, 60170615, 52644150, 264690, 264691, 33657023, 264692, 264693, 33657109, 33657182, 33657349, 18108370, 18108374, 18108377, 55811576, 35696423, 35695855, 264635, 264555, 264556, 56182323, 60170394, 264558, 264559, 83373044, 56526488, 87168518, 60432113, 22279002, 264482, 264563, 264484, 264567 |
| 2945 | 94233560 (5889, 5890) | Novel Protein sim. GBank gi 728831 sp P39188 ALU1_HUMAN - !!!!! ALU SUBFAMILY J WARNING ENTRY !!!!! | Contains protein domain (PF000096) - Zinc finger, C2H2 type | | UNCLASSIFIED | 60424179, 22278995, 22278996, 22278998, 22278999, 264259, 56182181, 29331824, 60424269, 60432289, 35696052, 264908, 265008, 60433356, 55812038, 264759, 55811386, 265018, 264681, 18108351, 264448, 264683, 264389, 264288, 264687, 56181562, 21906767, 21906768, 21906769, 35695917, 265020, 265021, 264693, 60431528, 55810784, 35696423, 35695855, 264630, 60170394, 83373044, 22279000, 264566, 264567 |

| | | | | | |
|------|-----------------------|---|--|--------------|---|
| 2946 | 94317315 (5891, 5892) | Novel Protein sim. GBank gij5441952 gb AAD43195.1 AF072864 peroxisomal membrane protein PMP 24 [Homo sapiens] | | UNCLASSIFIED | 264488, 264259, 264508, 264509, 264906, 264907, 264909, 264510, 264511, 265007, 264512, 264910, 264591, 264593, 18108351, 264764, 264288, 264684, 264769, 265021, 264692, 33657109, 264628, 264629, 18108374, 264631, 264634, 264636, 264637, 18108380, 264638, 264639, 83373044, 264555, 264566, 264486, 264567 |
| 2947 | 87362952 (5893, 5894) | Novel Protein sim. GBank gij3540281 gb AAC34383.1 - (AF056116) All-1 related protein [Fugu rubripes] | | UNCLASSIFIED | 22278995, 22278996, 22278997, 22278999, 29146498, 264508, 29331830, 265007, 265008, 265009, 60432228, 21906754, 265010, 265017, 265019, 264766, 264685, 21906765, 21906766, 21906767, 21906768, 21906769, 265020, 264628, 18108370, 264629, 264630, 18108387, 60432113 |
| 2948 | 87626527 (5895, 5896) | Novel Protein sim. GBank gij5566614 gb AAB65654.2 - (AF001533) mitogen-induced [Mus musculus] | | | 52646842, 22278995, 264259, 29331824, 29331825, 29331827, 29331830, 264909, 265007, 265009, 265019, 264763, 264684, 264288, 264685, 264688, 21906787, 264691, 264692, 264693, 18108374, 55811576, 18108385, 22279002, 264563, 264567 |
| 2949 | 88175545 (5897, 5898) | Novel Protein sim. GBank gij2132923 pir S67133 - probable membrane protein YOR240w - yeast (Saccharomyces cerevisiae) | | UNCLASSIFIED | 22278996, 22278997, 60432289, 29331826, 29331827, 29331828, 35696052, 29146499, 264104, 264107, 264905, 66712502, 264908, 60433356, 60433438, 87168559, 264764, 52644229, 56181562, 21906767, 21906768, 21906769, 265022, 60170615, 33657023, 35696423, 263981, 264558, 60432113, 22279002 |
| 2950 | 95086870 (5899, 5900) | Novel Protein sim. GBank gij466102 sp P34629 YOJ6_CAEEL - PUTATIVE AMINOPEPTIDASE ZK353.6 IN CHROMOSOME III | Contains protein domain (PF00863) - Cytosol aminopeptidase family | peptidase | 264488, 35696286, 264259, 35696052, 264907, 265007, 264910, 265017, 265018, 264288, 264768, 35695917, 265020, 18108362, 18108370, 18108379, 35696423, 65274791, 35695855, 264556, 56526486, 264486 |
| 2951 | 87392357 (5901, 5902) | Novel Protein sim. GBank gij4688902 emb CAB41450.1 - (AJ238248) centaurin beta2 [Homo sapiens] | | | 264683 |

| | | | | | |
|------|-----------------------|---|---|-------------------|--|
| 2952 | 95329852 (5903, 5904) | Novel Protein sim. GBank gjl5596683[embjCAB51405.1] - (AL096881) hypothetical protein [Homo sapiens] | Contains protein domain (PF00650) - CRAUTRIO domain. | transcript factor | 264687, 52645156, 21906766, 21906769, 22278996, 265020, 264690, 60432049, 264259, 264693, 29331822, 18108365, 29331825, 60432289, 33657109, 18108368, 29331827, 35696052, 27486262, 264508, 264905, 20281149, 264908, 264907, 29331830, 264908, 264909, 35695855, 264511, 265008, 265009, 264910, 264635, 264636, 60432229, 264638, 60433356, 264639, 264758, 87168518, 265017, 22279000, 22279002, 264760, 264563, 264482, 18108351, 264448, 264288, 18108396, 56994075, 22278896, 29331822, 29331824, 29331825, 29331827, 35696052, 264508, 264905, 264908, 264907, 264510, 264591, 264594, 33657402, 264595, 264596, 264758, 52646317, 21906754, 33657084, 52644296, 87168559, 264600, 264760, 264681, 18108351, 264764, 264369, 264288, 264687, 21906765, 21906766, 21906787, 21906768, 21906769, 35695917, 33657023, 18108384, 52645129, 33657109, 33657349, 18108374, 263978, 35695855, 264637, 264638, 87168518, 264482, 264563, 264565, 29331822 |
| 2953 | 88093575 (5905, 5906) | Novel Protein sim. GBank gjl119522[spIP10658]SERC, RABIT - PROBABLE PHOSPHOSERINE AMINOTRANSFERASE (PSAT) (ENDOMETRIAL PROGESTERONE-INDUCED PROTEIN) (EPIP) | Contains protein domain (PF00266) - Aminotransferases class-V | UNCLASSIFIED | 264687, 52645156, 21906766, 21906769, 22278996, 265020, 264690, 60432049, 264259, 264693, 29331822, 18108365, 29331825, 60432289, 33657109, 18108368, 29331827, 35696052, 27486262, 264508, 264905, 20281149, 264908, 264907, 29331830, 264908, 264909, 35695855, 264511, 265008, 265009, 264910, 264635, 264636, 60432229, 264638, 60433356, 264639, 264758, 87168518, 265017, 22279000, 22279002, 264760, 264563, 264482, 18108351, 264448, 264288, 18108396, 56994075, 22278896, 29331822, 29331824, 29331825, 29331827, 35696052, 264508, 264905, 264908, 264907, 264510, 264591, 264594, 33657402, 264595, 264596, 264758, 52646317, 21906754, 33657084, 52644296, 87168559, 264600, 264760, 264681, 18108351, 264764, 264369, 264288, 264687, 21906765, 21906766, 21906787, 21906768, 21906769, 35695917, 33657023, 18108384, 52645129, 33657109, 33657349, 18108374, 263978, 35695855, 264637, 264638, 87168518, 264482, 264563, 264565, 29331822 |
| 2954 | 88086288 (5907, 5908) | Novel Protein sim. GBank gjl4685261[refNP_005251.1]pGDF9 - growth differentiation factor 9 | Contains protein domain (PF00019) - Transforming growth factor beta like domain | tgf | 264687, 52645156, 21906766, 21906769, 22278996, 265020, 264690, 60432049, 264259, 264693, 29331822, 18108365, 29331825, 60432289, 33657109, 18108368, 29331827, 35696052, 27486262, 264508, 264905, 20281149, 264908, 264907, 29331830, 264908, 264909, 35695855, 264511, 265008, 265009, 264910, 264635, 264636, 60432229, 264638, 60433356, 264639, 264758, 87168518, 265017, 22279000, 22279002, 264760, 264563, 264482, 18108351, 264448, 264288, 18108396, 56994075, 22278896, 29331822, 29331824, 29331825, 29331827, 35696052, 264508, 264905, 264908, 264907, 264510, 264591, 264594, 33657402, 264595, 264596, 264758, 52646317, 21906754, 33657084, 52644296, 87168559, 264600, 264760, 264681, 18108351, 264764, 264369, 264288, 264687, 21906765, 21906766, 21906787, 21906768, 21906769, 35695917, 33657023, 18108384, 52645129, 33657109, 33657349, 18108374, 263978, 35695855, 264637, 264638, 87168518, 264482, 264563, 264565, 29331822 |
| 2955 | 87698426 (5909, 5910) | Novel Protein sim. GBank gjl3452473 (AF084205) - serine/threonine protein kinase TAO1 [Rattus norvegicus] | Contains protein domain (PF00069) - Eukaryotic protein kinase domain | kinase | 264687, 52645156, 21906766, 21906769, 22278996, 265020, 264690, 60432049, 264259, 264693, 29331822, 18108365, 29331825, 60432289, 33657109, 18108368, 29331827, 35696052, 27486262, 264508, 264905, 20281149, 264908, 264907, 29331830, 264908, 264909, 35695855, 264511, 265008, 265009, 264910, 264635, 264636, 60432229, 264638, 60433356, 264639, 264758, 87168518, 265017, 22279000, 22279002, 264760, 264563, 264482, 18108351, 264448, 264288, 18108396, 56994075, 22278896, 29331822, 29331824, 29331825, 29331827, 35696052, 264508, 264905, 264908, 264907, 264510, 264591, 264594, 33657402, 264595, 264596, 264758, 52646317, 21906754, 33657084, 52644296, 87168559, 264600, 264760, 264681, 18108351, 264764, 264369, 264288, 264687, 21906765, 21906766, 21906787, 21906768, 21906769, 35695917, 33657023, 18108384, 52645129, 33657109, 33657349, 18108374, 263978, 35695855, 264637, 264638, 87168518, 264482, 264563, 264565, 29331822 |
| 2956 | 85789743 (5911, 5912) | Novel Protein sim. GBank gjl4689254[gbAAD27830.1]AF12185 - (AF121857) sorting nexin 7 [Homo sapiens] | Contains protein domain (PF00787) - PX domain | | 264687, 52645156, 21906766, 21906769, 22278996, 265020, 264690, 60432049, 264259, 264693, 29331822, 18108365, 29331825, 60432289, 33657109, 18108368, 29331827, 35696052, 27486262, 264508, 264905, 20281149, 264908, 264907, 29331830, 264908, 264909, 35695855, 264511, 265008, 265009, 264910, 264635, 264636, 60432229, 264638, 60433356, 264639, 264758, 87168518, 265017, 22279000, 22279002, 264760, 264563, 264482, 18108351, 264448, 264288, 18108396, 56994075, 22278896, 29331822, 29331824, 29331825, 29331827, 35696052, 264508, 264905, 264908, 264907, 264510, 264591, 264594, 33657402, 264595, 264596, 264758, 52646317, 21906754, 33657084, 52644296, 87168559, 264600, 264760, 264681, 18108351, 264764, 264369, 264288, 264687, 21906765, 21906766, 21906787, 21906768, 21906769, 35695917, 33657023, 18108384, 52645129, 33657109, 33657349, 18108374, 263978, 35695855, 264637, 264638, 87168518, 264482, 264563, 264565, 29331822 |
| 2957 | 80933301 (5913, 5914) | Novel Protein sim. GBank gjl4503023[refNP_000089.1]pCPT2 - carnitine palmitoyltransferase II precursor | | cadherin | 264687, 52645156, 21906766, 21906769, 22278996, 265020, 264690, 60432049, 264259, 264693, 29331822, 18108365, 29331825, 60432289, 33657109, 18108368, 29331827, 35696052, 27486262, 264508, 264905, 20281149, 264908, 264907, 29331830, 264908, 264909, 35695855, 264511, 265008, 265009, 264910, 264635, 264636, 60432229, 264638, 60433356, 264639, 264758, 87168518, 265017, 22279000, 22279002, 264760, 264563, 264482, 18108351, 264448, 264288, 18108396, 56994075, 22278896, 29331822, 29331824, 29331825, 29331827, 35696052, 264508, 264905, 264908, 264907, 264510, 264591, 264594, 33657402, 264595, 264596, 264758, 52646317, 21906754, 33657084, 52644296, 87168559, 264600, 264760, 264681, 18108351, 264764, 264369, 264288, 264687, 21906765, 21906766, 21906787, 21906768, 21906769, 35695917, 33657023, 18108384, 52645129, 33657109, 33657349, 18108374, 263978, 35695855, 264637, 264638, 87168518, 264482, 264563, 264565, 29331822 |
| 2958 | 87440014 (5915, 5916) | Novel Protein sim. GBank gjl4240257[dbjBAA74907.1] - (AB020691) KIAA0884 protein [Homo sapiens] | | | 264687, 52645156, 21906766, 21906769, 22278996, 265020, 264690, 60432049, 264259, 264693, 29331822, 18108365, 29331825, 60432289, 33657109, 18108368, 29331827, 35696052, 27486262, 264508, 264905, 20281149, 264908, 264907, 29331830, 264908, 264909, 35695855, 264511, 265008, 265009, 264910, 264635, 264636, 60432229, 264638, 60433356, 264639, 264758, 87168518, 265017, 22279000, 22279002, 264760, 264563, 264482, 18108351, 264448, 264288, 18108396, 56994075, 22278896, 29331822, 29331824, 29331825, 29331827, 35696052, 264508, 264905, 264908, 264907, 264510, 264591, 264594, 33657402, 264595, 264596, 264758, 52646317, 21906754, 33657084, 52644296, 87168559, 264600, 264760, 264681, 18108351, 264764, 264369, 264288, 264687, 21906765, 21906766, 21906787, 21906768, 21906769, 35695917, 33657023, 18108384, 52645129, 33657109, 33657349, 18108374, 263978, 35695855, 264637, 264638, 87168518, 264482, 264563, 264565, 29331822 |

| | | | | | |
|------|-----------------------|---|---|------------------|--|
| 2859 | 95109420 (5917, 5918) | Novel Protein sim. GBank gj1988221 (U33005) - Tbc1 [Mus musculus] | Contains protein domain (PF00566) - TBC domain | oncogene | 263994, 22278997, 264259, 60432049, 29331826, 29331828, 35696052, 29331830, 66712502, 56182435, 265006, 264512, 265008, 265009, 60433356, 60433438, 264596, 265017, 265018, 264683, 264288, 264766, 264768, 21906766, 21906767, 21906769, 265020, 60170615, 264692, 27486265, 18108374, 65274791, 35695855, 83373044, 56526486, 60432113 |
| 2860 | 87420091 (5918, 5920) | | | UNCLASSIFIED | 35696286, 56182435, 87168474, 265010, 60170615, 35698423, 56182323, 18108383, 87168518, 264483 |
| 2861 | 95413416 (5921, 5922) | Novel Protein sim. GBank gj1596646[emb]CAB05177.2] - (Z82266) predicted using GeneFinder; similar to WD domain; G-beta repeats [Caenorhabditis elegans] | Contains protein domain (PF00400) - WD domain, G-beta repeat | transcriptfactor | 22278997, 22278999, 264259, 29331822, 29331824, 29331826, 29331828, 264907, 264908, 52644045, 265006, 33657402, 21906754, 87168474, 265011, 87168559, 265017, 21906769, 265020, 60170615, 264692, 33657023, 35695763, 18108370, 18108374, 35698423, 264632, 264636, 18108385, 87168518, 22279002, 264564, 264567 |
| 2862 | 87912700 (5923, 5924) | | | UNCLASSIFIED | 35696286, 22278997, 264092, 264094, 264259, 29331824, 66714117, 29331825, 60432289, 29331826, 29331827, 29331828, 35696052, 264508, 264905, 264509, 264907, 264908, 264909, 264510, 264512, 264593, 264594, 60433438, 264758, 52646317, 264602, 264603, 264605, 264760, 264762, 264764, 264288, 264766, 264686, 264768, 264769, 35695917, 265020, 264691, 264634, 264636, 264637, 264638, 264639, 18108385, 264563, 264565, 264566, 264567, 264486 |
| 2863 | 95313464 (5925, 5926) | Novel Protein sim. GBank gj14240223[dj]BAA74890.1] - (AB020874) KIAA0867 protein [Homo sapiens] | Contains protein domain (PF00010) - Helix-loop-helix DNA-binding domain | transcriptfactor | 18108392, 56994075, 22278998, 22278999, 29331822, 29331825, 29331826, 29331827, 29331828, 265007, 265008, 264592, 264594, 21906754, 265016, 264760, 264687, 29148627, 29148784, 265020, 33657023, 264693, 65274620, 33657182, 27486261, 264629, 55810764, 35698423, 264555, 264636, 264637, 264557, 264558, 264563 |
| 2864 | 94324617 (5927, 5928) | | | UNCLASSIFIED | 264259, 29331828, 33657402, 265017, 265018, 264692, 18108368, 35698423, 83373044, 18108388 |

| | | | | | | |
|------|-----------------------|--|---|--|-------------------|---|
| 2865 | 80384762 (5929, 5930) | Novel Protein sim. GBank gi4885447 rel NP_005452.1 pKRML - Kreisler (mouse) mal- related leucine zipper homolog | | | transcript factor | 264259, 29331826, 264508, 264509, 264905, 264907, 264908, 264909, 264511, 265008, 264910, 264591, 264593, 264594, 33657402, 265011, 264760, 264762, 264764, 264288, 264685, 264766, 264692, 33657109, 264628, 264629, 35695855, 264630, 264631, 264632, 264634, 264635, 264636, 264637, 264638, 264639, 264563, 264567, 18108391, 60432289, 264682, 264448 |
| 2866 | 91725248 (5931, 5932) | Novel Protein sim. GBank gi5262751 emb CAB45690.1 - (AJ243177) Xenopus RPA interacting protein alpha [Xenopus laevis] | | | | |
| 2867 | 94658303 (5933, 5934) | Novel Protein sim. GBank gi524225 (U19181) - Rabin3 [Rattus norvegicus] | | | UNCLASSIFIED | 264488, 264508, 264509, 264908, 264909, 264511, 264910, 264594, 264758, 85658542, 264762, 264764, 265021, 264556, 18108381, 264564, 264486 |
| 2868 | 95302778 (5935, 5936) | Novel Protein sim. GBank gi4929715 gb AAD34118.1 AF15188 - (AF151881) CGI-123 protein [Homo sapiens] | Contains protein domain (PF00097) - Zinc finger, C3HC4 type (RING finger) | | | 264687, 52845156, 21906765, 52846365, 21906767, 18108398, 35696423, 22278996, 3589286, 22278997, 265020, 22278999, 265021, 265022, 264093, 264636, 264690, 52844150, 264259, 33657023, 52845080, 264693, 29331822, 56182181, 29331824, 66714117, 29331825, 33109954, 52645129, 29331826, 21906754, 33657182, 29331827, 29331828, 35696052, 27486262, 87168518, 87168474, 265010, 87168559, 265018, 22278000, 265019, 22279002, 264563, 18108351, 264906, 264907, 264448, 68712502, 264566, 264369, 264288 |
| 2869 | 95310957 (5937, 5938) | Novel Protein sim. GBank gi3024743 sp O24734 THSA_SULS7 - THERMOSOME, ALPHA SUBUNIT (CHAPERONIN ALPHA SUBUNIT) | | | eph | 52646842, 22278996, 22278998, 22278999, 60432049, 264259, 29331824, 29331825, 29331826, 29331828, 264509, 264909, 52644045, 56182435, 265009, 60433438, 55812038, 21906754, 265011, 87168559, 265018, 265019, 264448, 264288, 264369, 52644229, 21906766, 21906768, 21906769, 29148784, 265020, 265021, 52644150, 264691, 33657109, 18108374, 56182323, 60170394, 87168518, 60432113, 22279000 |
| 2870 | 88088071 (5939, 5940) | Novel Protein sim. GBank gi3165407 (AC004755) - fos37502.1 [Homo sapiens] | Contains protein domain (PF00046) - Homeobox domain | | homeobox | |

| | | | | | |
|------|-----------------------|---|---|--------------|--|
| 2971 | 94186830 (5941, 5942) | Novel Protein sim. GBank gi 728837 sp P39194 ALU7_HUMAN - IIII ALU SUBFAMILY SQ WARNING ENTRY IIII | | Im7 | 264488, 56182575, 35696286, 56994075, 29331824, 29331826, 29146489, 264508, 264905, 264907, 264112, 264910, 21906754, 87168559, 265018, 265019, 18108351, 264689, 21908765, 21906767, 21906768, 265020, 265021, 60170615, 18108364, 264628, 264629, 18108374, 264636, 264556, 264558, 83373044, 18108384, 18108385, 87168518, 264584, 264587 |
| 2972 | 86625943 (5943, 5944) | Novel Protein sim. GBank gi 728838 sp P39193 ALU6_HUMAN - IIII ALU SUBFAMILY SP WARNING ENTRY IIII | | kinase | 265017, 35695917, 265021, 33657109, 22279002, 264563 |
| 2973 | 91215301 (5945, 5946) | Novel Protein sim. GBank gi 2746789 (AF040842) - No definition line found [Caenorhabditis elegans] | | UNCLASSIFIED | 29331822, 264692, 33657349, 55811576, 264563 |
| 2974 | 91673002 (5947, 5948) | Novel Protein sim. GBank gi 786117 (L41834) - nuclear protein [Ensis minor] | | UNCLASSIFIED | 18108392, 52644507, 56182575, 56181686, 22278995, 22278996, 35696286, 22278997, 22278998, 22278999, 264259, 52645080, 29331824, 29331825, 66714117, 60424269, 29331826, 29331827, 29331828, 35696052, 66712502, 284908, 52644045, 265007, 264910, 265009, 60433438, 33109954, 21906754, 55811386, 52644296, 87168474, 87168559, 265017, 265018, 265019, 18108351, 264448, 264369, 264288, 52644229, 18108358, 21906765, 21906767, 21906768, 35695917, 265020, 265021, 52644150, 264691, 264692, 33657023, 27486262, 27486284, 35695763, 18108370, 18108376, 55810764, 55811576, 35696423, 35695855, 264630, 264635, 264557, 52844332, 264558, 83373044, 18108387, 87168518, 60432113, 22278000, 264482, 264487 |
| 2975 | 95325213 (5949, 5950) | Novel Protein sim. GBank gi 3880812 emb CAA19508 - (AL023639) similar to HECT-domain (ubiquitin-transferase); cDNA EST YK480d10.5 comes from this gene [Caenorhabditis elegans] | Contains protein domain (PF00632) - HECT-domain (ubiquitin- transferase). | ubiquitin | 29331824, 29331827, 29331828, 264910, 85858542, 265011, 265018, 264448, 264288, 264769, 21906767, 265020, 264691, 264559, 83373044 |
| 2976 | 8771202 (5951, 5952) | Novel Protein sim. GBank gi 5678136 gb AA046874.1 AF16093 - (AF160934) BcDNA.LD14189 [Drosophila melanogaster] | | transport | 22278996, 264906, 265007, 265010, 265011, 265017, 265019, 18108351, 264685, 264689, 18108370, 264639, 18108385 |
| 2977 | 91725254 (5953, 5954) | Novel Protein sim. GBank gi 5262751 emb CAB45690.1 - (AJ243177) Xenopus RPA interacting protein alpha [Xenopus laevis] | | UNCLASSIFIED | 264509, 264288 |

| | | | | | |
|------|-----------------------|---|--|--------------|--|
| 2878 | 87332059 (5955, 5956) | Novel Protein sim. GBank gi 746549 (U23522) - No definition line found [Caenorhabditis elegans] | Contains protein domain (PF00480) - ROK family | UNCLASSIFIED | 22278995, 22278996, 22278997, 22278999, 264259, 60432289, 29331827, 29146499, 56182435, 265008, 265007, 265009, 60433356, 60433438, 21906754, 265010, 265011, 265017, 265018, 265019, 264288, 264685, 264688, 21906765, 21906766, 21906767, 21906768, 21906769, 265020, 265021, 265022, 35696423, 264639, 60432113, 22279000, 22279002 |
| 2879 | 91725256 (5957, 5958) | Novel Protein sim. GBank gi 5262751 emb CAB45690.1 - (AJ243177) Xenopus RPA interacting protein alpha [Xenopus laevis] | complement | | 264488, 65274572, 56994075, 22278999, 264093, 29331822, 29331824, 264288, 55811957, 33657023, 33657109, 18108370, 55811576, 56182323, 60432113, 264482, 265009, 21906767, 263981, 22279000 |
| 2880 | 86296600 (5959, 5960) | | | UNCLASSIFIED | 264629, 264564 |
| 2881 | 87376330 (5961, 5962) | | | | |
| 2882 | 85303675 (5963, 5964) | Novel Protein sim. GBank gi 4929767 gb AAD34144.1 AF15190 - (AF151907) CGI-149 protein [Homo sapiens] | | | 22278995, 56994075, 22278996, 22278997, 22278998, 22278999, 264092, 29331824, 29331827, 29331828, 264905, 264591, 264592, 264594, 264595, 264596, 33657084, 264448, 21906765, 21906766, 21906767, 21906768, 21906769, 265020, 265022, 18108365, 33657182, 33657349, 35696423, 83373044, 22279000, 22279002 |
| 2883 | 91725256 (5965, 5966) | Novel Protein sim. GBank gi 5262751 emb CAB45690.1 - (AJ243177) Xenopus RPA interacting protein alpha [Xenopus laevis] | | | 60424179, 52646842, 18108398, 22278997, 264093, 60432049, 264259, 29331822, 60432289, 33656970, 264905, 52644045, 265006, 60431735, 87168474, 265018, 265019, 18108351, 264448, 21906765, 21906768, 35695917, 33657023, 52645129, 18108370, 35696423, 83373044, 56526486, 60432113, 264404, 22279002 |
| 2884 | 84136467 (5967, 5968) | Novel Protein sim. GBank gi 2393734 (AC002542) - similar to C. elegans F11A10.5; 80% similarity to Z68297 (PID:g1130619) [Homo sapiens] | ATPase_associated | | |
| 2885 | 87099072 (5969, 5970) | Novel Protein sim. GBank gi 103160 pir S22126 - finger protein unkempt - fruit fly (Drosophila melanogaster) | UNCLASSIFIED | | 264910, 55812038, 56181562, 55811957, 264628, 55810764, 264632, 264635, 60432113 |
| 2886 | 86284861 (5971, 5972) | | | | 55811957, 264566 |
| 2887 | 86455934 (5973, 5974) | | UNCLASSIFIED | | 264369 |

| | | | | | |
|------|-----------------------|---|--|--------------|--|
| 2988 | 95357753 (5975, 5976) | Novel Protein sim. GBank gii4678028[gb AAD27002.1] - (AF077207) HSPC021 [Homo sapiens] | | UNCLASSIFIED | 264488, 65274572, 22278995, 22278996, 22278997, 22278999, 264092, 264094, 264259, 60432049, 29331824, 29331828, 60432289, 35696052, 29331828, 264107, 264905, 264907, 264908, 66712502, 264828, 264909, 56182435, 265008, 265007, 265008, 60170831, 60432229, 264593, 60433356, 264757, 60433438, 21906754, 265010, 265011, 87168559, 265017, 265018, 264682, 264448, 264389, 264288, 264685, 52644229, 21906765, 21906767, 21906769, 35895917, 265021, 265022, 52644150, 264690, 33657023, 65274620, 263967, 33657109, 27486262, 18108370, 18108372, 18108374, 55810764, 65274791, 35695855, 264635, 264636, 264637, 263981, 264638, 56182323, 83373044, 60432113, 22279000, 264563, 264564, 264565, 264566, 264567 |
| 2989 | 81225118 (5977, 5978) | Novel Protein sim. GBank gii113671[sp P23964 ALUF_HUMAN - IIII ALU CLASS F WARNING ENTRY IIII] | | | kinase |
| 2990 | 87330444 (5979, 5980) | Novel Protein sim. GBank gii2829836[sp P97346 RHOD_MOUSE - RHO-RELATED GTP-BINDING PROTEIN RHOD] | Contains protein domain (PF00071) - Ras family | | oncogene |
| 2991 | 94325361 (5981, 5982) | | | | |
| 2992 | 85425164 (5983, 5984) | | | | |
| 2993 | 94325363 (5985, 5986) | | | | |
| 2994 | 94136634 (5987, 5988) | Novel Protein sim. GBank gii249654[sp Q50658 YU02_MYCTU - HYPOTHETICAL 29.7 KD PROTEIN CY339.02] | | UNCLASSIFIED | 264563 |
| | | | | UNCLASSIFIED | 264259, 265019, 264689, 18108385 |
| | | | | UNCLASSIFIED | 264488, 29331822, 265017, 264761, 21906769, 65274791, 263981, 264565 |
| | | | | transport | 22278994, 22278995, 56994075, 22278997, 22278999, 264259, 29331822, 29331824, 29331825, 29331826, 29331827, 29331828, 265006, 265009, 264910, 33109954, 87168474, 87168559, 265018, 265019, 264448, 264288, 21906766, 21906767, 21908768, 21906769, 265021, 265022, 33657023, 264693, 35695855, 83373044, 18108385, 22279000, 264565, 264566 |
| 2995 | 87591070 (5989, 5990) | Novel Protein sim. GBank gii2734081 (AF000195) - similar to oxysterol-binding proteins [Caenorhabditis elegans] | | | |
| 2996 | 91013798 (5991, 5992) | Novel Protein sim. GBank gii2829912 (AC002291) - Similar ATP-dependent RNA Helicase [Arabidopsis thaliana] | Contains protein domain (PF00270) - DEAD/DEAH box helicase | | helicase |
| | | | | | 65274572, 35696286, 264259, 29331824, 35696052, 29148499, 264508, 264907, 265007, 265008, 60433438, 18108348, 265017, 264681, 264683, 264288, 264766, 264769, 264689, 35695917, 60170615, 33657023, 264692, 264634, 264555, 18108381, 18108382, 18108388, 264484 |

| | | | | | |
|------|-----------------------|---|--|--------------|--|
| 2997 | 87627440 (5993, 5994) | Novel Protein sim. GBank gi 4589652 dbj BAA76848.1 - (AB023221) KIAA1004 protein [Homo sapiens] | | homeobox | 264488, 56182575, 264259, 66714117, 29331826, 35696052, 264508, 264509, 264907, 264908, 265006, 87168474, 265019, 264448, 264682, 264685, 264768, 21908764, 21908766, 21908768, 21908769, 27486261, 18108374, 35696423, 264634, 264635, 264636, 264557, 18108385, 87168518 |
| 2998 | 88095381 (5995, 5996) | Novel Protein sim. GBank gi 3947589 emb CAA22252 - (AL034364) cDNA EST yk255b9.3 comes from this gene; cDNA EST yk255b9.5 comes from this gene; cDNA EST EMBL: M75923 comes from this gene [Caenorhabditis elegans] | | UNCLASSIFIED | 52646365, 22278997, 264508, 264908, 18108351, 21908765, 21908767, 18108370, 18108374, 35696423, 264636, 264639 |
| 2999 | 94847055 (5997, 5998) | Novel Protein sim. GBank gi 15408 sp P18835 CC19_CAEL - CUTICLE COLLAGEN 19 | | UNCLASSIFIED | 56182575, 22278996, 29147620, 29331825, 29146498, 29146499, 264905, 66712502, 265006, 265009, 21908754, 85658542, 18108351, 29148627, 29148629, 60170615, 33657109, 27486262, 18108370, 18108374, 264556, 264557, 264558, 60170394, 18108385, 264563 |
| 3000 | 95099370 (5999, 6000) | Novel Protein sim. GBank gi 1163174 (U32575) - similar to yeast Sec6p. Swiss-Prot Accession Number P32844; similar to mammalian B94. Swiss-Prot Accession Number Q03169; Method: conceptual translation supplied by author [Rattus norvegicus] | | UNCLASSIFIED | 264887, 22278997, 22278999, 264259, 29331822, 29331824, 35696052, 29146498, 264508, 264905, 264908, 264907, 264908, 264909, 264510, 264511, 265006, 265007, 265008, 265009, 264910, 33657402, 264757, 264595, 264596, 264758, 21908754, 265011, 264600, 265017, 265018, 264605, 265019, 264760, 264761, 264762, 264681, 264682, 264764, 264288, 264685, 264766, 264686, 264768, 264769, 21908765, 21908768, 35695917, 265020, 264691, 264692, 33657023, 264693, 33657109, 33657182, 27486261, 264628, 264629, 18108374, 18108376, 35696423, 35695855, 264630, 264631, 264632, 264634, 264635, 264636, 264637, 264556, 264638, 264639, 60170394, 83373044, 20798451, 22279002, 264563, 264486, 264567 |
| 3001 | 86078454 (6001, 6002) | Novel Protein sim. GBank gi 2078470 (AC002073) - Putative gene. Genscan predictions confirmed by EST splicing.; coded for by human cDNAs AA122029 (NID:g1678048), D31562 (NID:g644442), AA158721 (NID:g1733515), R59840 (NID:g830335) and F13082 (NID:g709111) [Homo sapiens] | | calthepsin | 18108394, 52646842, 56182575, 29331824, 29331825, 29331827, 264910, 33109954, 52644296, 265017, 265019, 264288, 265020, 265021, 52644150, 264692, 35695763, 55810764, 35696423, 56182323, 18108387, 264563, 264564 |
| 3002 | 87718167 (6003, 6004) | Novel Protein sim. GBank gi 3599478 (AF085185) - Myosin-1A [Acanthamoeba castellanii] | | UNCLASSIFIED | 264488, 29331824, 29331825, 29331826, 29331827, 29331828, 264908, 264510, 265009, 21908754, 264682, 264688, 33657023, 264565 |

| | | | | | |
|------|-----------------------|---|--|-------------------|---|
| 3003 | 86648078 (6005, 6006) | Novel Protein sim. GBank gij1754969 (U30292) - collagen type XIII alpha-1 chain [Mus musculus] | Contains protein domain (PF01391) - Collagen triple helix repeat (20 copies) | collagen | 264512, 264593, 264564, 264567, 264486 |
| 3004 | 88066876 (6007, 6008) | Novel Protein sim. GBank gij2224629jdbj(BAA20802) - (AB002342) KIAA0344 [Homo sapiens] | Contains protein domain (PF01360) - Monooxygenase | oxygenase | 29331830, 21906769, 264691, 33657109, 263972, 18108385 |
| 3005 | 87794843 (6009, 6010) | Novel Protein sim. GBank gij4680659jgb/AAD27719.1(A13294) - (AF132944) CGI-10 protein [Homo sapiens] | Contains protein domain (PF00023) - MHC | MHC | 29331822, 29331824, 29331827, 60433438, 265011, 265019, 21906766, 21908767, 21908768, 265020, 33657023, 33657349, 60170394, 22279002, 264567 |
| 3006 | 87422224 (6011, 6012) | Novel Protein sim. GBank gij3930525 (AF064447) - sex-determination protein homolog Fem1a [Mus musculus] | Contains protein domain (PF00096) - Zinc finger, C2H2 type | transcript factor | 264259, 29331822, 264512, 21906754, 265018, 264687, 21906765, 264691, 264555, 264556, 264558, 18108385 |
| 3007 | 80936005 (6013, 6014) | Novel Protein sim. GBank gij2565052 (U80738) - CAGH1a [Homo sapiens] | Contains protein domain (PF00501) - AMP-binding enzyme | synthase | 52644507, 52645156, 65274572, 264809, 264512, 265018, 264760, 264448, 264765, 264889, 60170615, 18108374, 20281152, 264636, 52644332 |
| 3008 | 80416249 (6015, 6016) | Novel Protein sim. GBank gij3127193 (AF062389) - kidney-specific protein [Rattus norvegicus] | Contains protein domain (PF01923) - Protein of unknown function | UNCLASSIFIED | 264805, 264593, 264766, 264636 |
| 3009 | 91213387 (6017, 6018) | Novel Protein sim. GBank gij4927370jgb/AAD33084.1(A13294) - (AF067972) DNA cytosine methyltransferase 3 alpha [Homo sapiens] | Contains protein domain (PF00153) - Mitochondrial carrier proteins | transport | 52646842, 58182575, 22278995, 22278996, 264259, 29331825, 29331826, 29331827, 29331828, 33696052, 264508, 264509, 264907, 58182435, 264511, 265007, 264512, 265008, 264757, 264758, 5812038, 264759, 33109954, 21906754, 265010, 265011, 264600, 265017, 265018, 265019, 264760, 18108351, 264288, 264369, 21906764, 21906765, 21906767, 5811957, 265020, 265021, 264691, 18108368, 27486262, 20281149, 18108370, 5811578, 264637, 264556, 264557, 18108381, 264558, 58182323, 264559, 18108385, 18108388, 22279002, 264486 |
| 3010 | 95317217 (6019, 6020) | Novel Protein sim. GBank gij4927370jgb/AAD33084.1(A13294) - (AF067972) DNA cytosine methyltransferase 3 alpha [Homo sapiens] | Contains protein domain (PF00153) - Mitochondrial carrier proteins | transport | 264686, 264687, 21906767, 21906769, 5811957, 22278995, 35695917, 22278996, 22278997, 265020, 265021, 60170615, 264692, 33657023, 29331822, 264693, 18108364, 29331824, 33657109, 60432289, 29331827, 27486261, 29331828, 264508, 264909, 5811576, 35695855, 265008, 264556, 60433438, 83373044, 18108387, 65274727, 60432113, 265017, 22279000, 265019, 264564, 264682, 264764 |
| 3011 | 94323597 (6021, 6022) | Novel Protein sim. GBank gij5052319jgb/AAD38501.1(A11883) - (AF118838) citrin; adult-onset type II citrullinemia protein [Homo sapiens] | Contains protein domain (PF00153) - Mitochondrial carrier proteins | transport | 35696052, 58182435, 264758, 21906754, 265018, 264760, 264762, 18108351, 264682, 264448, 21906766, 65274620, 18108374, 264482, 264564 |
| 3012 | 87753087 (6023, 6024) | | UNCLASSIFIED | UNCLASSIFIED | 263972 |

| | | | | | |
|------|-----------------------|--|--|-------------------|--|
| 3013 | 81238768 (6025, 6026) | Novel Protein sim. GBank gi 3702286 (AC005787) - R33374_1 [Homo sapiens] | Contains protein domain (PF00400) - WD domain, G-beta repeat | transcriptfactor | 264488, 263994, 35696286, 22278997, 264259, 29331824, 60424269, 66714117, 35696052, 264905, 264908, 264907, 264908, 264909, 56182435, 264511, 264512, 264910, 264591, 264592, 264593, 264594, 33657402, 60433438, 264595, 264596, 55812038, 264758, 33109954, 21906754, 265010, 265018, 264604, 264760, 264682, 264683, 264764, 264369, 264288, 264765, 264766, 264686, 264768, 264687, 21906767, 35695917, 265020, 33657023, 264692, 264693, 33657109, 264628, 264629, 55811576, 35696423, 35695855, 264630, 264631, 264632, 264634, 264635, 264636, 264637, 264638, 264639, 83373044, 264563, 264565, 264566, 264587, 264760 |
| 3014 | 79877263 (6027, 6028) | Novel Protein sim. GBank gi 3878374 emb CAA93081 (Z68879) Similarity to Yeast Chl12p protein (PIR Acc. No. S54453); cDNA EST EMBL:D27950 comes from this gene; cDNA EST EMBL:D27948 comes from this gene; cDNA EST EMBL:D33447 comes from this gene; cDNA EST EMBL:D33316 comes from... | | ATPase_associated | |
| 3015 | 86995466 (6029, 6030) | | | | 22278995, 22278996, 22278997, 264259, 29331824, 29331828, 264908, 265007, 265008, 264910, 265011, 265017, 265019, 264691, 33657109, 18108370, 35695855, 264556, 264564 |
| 3016 | 87759945 (6031, 6032) | Novel Protein sim. GBank gi 168819 sp P41733 CC91_YEAST - CELL DIVISION CONTROL PROTEIN 91 | | UNCLASSIFIED | 52644507, 52646842, 56994075, 52645080, 29331822, 29331824, 35696052, 33656970, 52644045, 264596, 33657084, 265017, 265019, 52644229, 21906767, 35695917, 52644150, 33657023, 33657109, 27486261, 27486262, 27486264, 33657349, 27486265, 35695763, 35695855, 87168518 |
| 3017 | 95011154 (6033, 6034) | Novel Protein sim. GBank gi 4569658 dbj BAA76851.1 (AB023224) KIAA1007 protein [Homo sapiens] | | | 264488, 18108397, 22278996, 35696286, 22278999, 264259, 29331822, 60432289, 264908, 29331830, 264909, 56182435, 265006, 265007, 265008, 265009, 264591, 60433358, 60433438, 52646317, 21906754, 55811386, 265010, 265011, 87168559, 265017, 265018, 265019, 264288, 264687, 21906765, 21906766, 21906767, 21906769, 265020, 265022, 65274820, 52645129, 33657109, 33657182, 18108370, 263972, 18108374, 264631, 52644332, 83373044, 18108385, 18108388, 56526486, 87168518, 264404, 60432113, 22279000, 264567 |

| | | | | | |
|------|-----------------------|---|---|-------------------|--|
| 3018 | 11073891 (6035, 6036) | Novel Protein sim. GBank gjl3219332 (AC004020) - Unknown gene product [Homo sapiens] | | oncogene | 264558 264569, 52644507, 18108394, 65274572, 56182575, 22278994, 22278995, 56994075, 22278998, 22278999, 264259, 29331822, 29331824, 60432289, 29331827, 264908, 56182435, 265007, 265008, 60432229, 264593, 60433356, 55812038, 21908754, 87168474, 265011, 87168559, 265017, 265018, 265019, 264681, 18108351, 264448, 264682, 264683, 18108354, 264685, 264687, 264689, 21908766, 21908768, 21908769, 52644150, 264690, 264691, 33657023, 264692, 264693, 33657109, 52645129, 33657349, 264629, 65274791, 264634, 52644332, 56182323, 18108385, 87168518, 22279000, 22279002, 284563 |
| 3020 | 94318251 (6039, 6040) | Novel Protein sim. GBank gjl3414809 (AF061529) - rjs [Mus musculus] | Contains protein domain (PF00415) - Regulator of chromosome condensation (RCC1) | ATPase-associated | 264488, 263994, 35696286, 264259, 264508, 264905, 264508, 264906, 264907, 264908, 264909, 264510, 264910, 60174639, 264600, 264603, 264760, 264762, 264682, 264763, 264764, 264288, 264369, 264766, 264687, 264688, 264769, 55811957, 35695917, 33657023, 284628, 35696423, 35695855, 264630, 264632, 264634, 284635, 264636, 264637, 264556, 264557, 264638, 264639, 83373044, 18108385, 264564, 264567, 264486 |
| 3021 | 80478512 (6041, 6042) | Novel Protein sim. GBank gjl3880889 (emb) [CAB09005] - (Z95559) cDNA EST yk236d4.5 comes from this gene; cDNA EST EMBL: C13455 comes from this gene; cDNA EST yk329g6.5 comes from this gene; cDNA EST CEMSH45R comes from this gene [Caenorhabditis elegans] | | | 264769, 264629, 264482 |
| 3022 | 87718500 (6043, 6044) | | | UNCLASSIFIED | 264259, 29331826, 29331828, 264288, 264566 |
| 3023 | 95305484 (6045, 6046) | Novel Protein sim. GBank gjl416592 (spIP32323JAGA1_YEAST - A-AGGLUTININ ATTACHMENT SUBUNIT PRECURSOR | Contains protein domain (PF00614) - Phospholipase D. Active site motif | UNCLASSIFIED | 264498, 22278995, 35696286, 22278997, 29331826, 35696052, 264907, 29331830, 52644045, 56182435, 60432229, 264592, 60433356, 60433438, 264689, 21908767, 55811957, 35695917, 265021, 18108376, 263978, 264635, 264558, 22279000 |
| 3024 | 86675305 (6047, 6048) | | | UNCLASSIFIED | 60432049, 264760, 21908769, 55811957, 35695917, 264690, 264555, 264559 |
| 3025 | 65706629 (6049, 6050) | Novel Protein sim. GBank gjl295671 (L11275) - selected as a weak suppressor of a mutant of the subunit AC40 of DNA dependant RNA polymerase I and III [Saccharomyces cerevisiae] | | | 264593, 55811576 |

| | | | | | |
|------|-----------------------|--|---|--------------|--|
| 3026 | 87643662 (6051, 6052) | Novel Protein sim. GBank gi 3024052 sp P97924 KARI_RAT - KALIRIN (PAM COOH-TERMINAL INTERACTOR PROTEIN 10) (P-CIP10) | | UNCLASSIFIED | 22278996, 22278997, 264490, 29331825, 264111, 265007, 60170831, 265010, 87188559, 265019, 21906765, 29148627, 263967, 20281148, 20281069, 263975, 263977, 20281071, 56526486, 22279000 |
| 3027 | 94844563 (6053, 6054) | Novel Protein sim. GBank gi 4929647 gb AAD34084.1 AF15184 - (AF151847) CGI-89 protein [Homo sapiens] | Contains protein domain (PF01529) - DHC zinc finger domain | UNCLASSIFIED | 18108394, 22278995, 22278996, 35696286, 22278997, 22278998, 60432048, 264259, 29331822, 29331824, 66714117, 29331825, 60432289, 29331826, 29331827, 29331828, 35696052, 264905, 264907, 29331830, 264908, 264909, 264510, 265007, 265008, 265009, 264910, 33657402, 264596, 21906754, 265010, 265011, 87188559, 264600, 265018, 18108351, 264682, 264683, 264764, 264288, 264685, 264687, 264769, 264689, 21906765, 21906766, 21906767, 21906768, 21906769, 29148629, 35695917, 265020, 265021, 265022, 52644150, 264692, 33657023, 264693, 52645129, 33657109, 27486261, 18108374, 55811576, 35696423, 65274791, 264636, 264556, 264557, 264638, 60170394, 264639, 264558, 83373044, 18108385, 56526486, 22279000, 22279002 |
| 3028 | 94231987 (6055, 6056) | Novel Protein sim. GBank gi 3080521 emb CAA18650 - (AL022599) hypothetical protein [Schizosaccharomyces pombe] | Contains protein domain (PF00400) - WD domain, G-beta repeat | UNCLASSIFIED | 22278995, 22278996, 22278997, 22278999, 264259, 29331824, 29331827, 35696052, 264908, 265007, 265008, 265009, 60170831, 21906754, 265011, 87188559, 265018, 264762, 264683, 264765, 264689, 21906765, 21906768, 21906769, 29148629, 35695917, 265021, 265022, 33657109, 27486265, 264628, 264629, 18108374, 35696423, 35695855, 264638, 60170394, 22279000, 22279002, 264482, 264564 |
| 3029 | 87619284 (6057, 6058) | | | UNCLASSIFIED | 22278997, 22278999, 29331827, 264905, 264509, 264909, 264510, 264511, 264512, 87168474, 265019, 18108351, 21906768, 264534, 264690, 264693, 263969, 18108370, 264558, 22279000, 22279002, 264482 |
| 3030 | 87544828 (6059, 6060) | Novel Protein sim. GBank gi 3757726 emb CAA18782 - (AL022727) dJ8019.1 (olfactory receptor-like protein [Hs6M1-1]) [Homo sapiens] | Contains protein domain (PF00001) - Im7 7 transmembrane receptor (rhodopsin family) | im7 | |
| 3031 | 91677853 (6061, 6062) | Novel Protein sim. GBank gi 4530587 gb AAD22105.1 - (AF132000) TADA1 protein [Homo sapiens] | | UNCLASSIFIED | 22278995, 22278997, 22278998, 22278999, 264259, 29331822, 29331824, 29331826, 29331827, 29331828, 35696052, 264908, 265018, 21906765, 21906766, 21906767, 21906768, 265021, 263974, 18108374, 264558, 56526486, 22279000, 22279002 |

| | | | | | |
|------|-----------------------|--|---|--------------|--|
| 3032 | 94130124 (6083, 6084) | Novel Protein sim. GBank gij101951 (U37429) - similar to M. musculus MER5 and other AHP/CTSA proteins [Caenorhabditis elegans] | Contains protein domain (PF00534) - Glycosyl transferases group 1 | synthase | 22278996, 35696286, 264259, 29331824, 29331828, 264907, 29331830, 264758, 33109954, 87168474, 87168559, 265019, 264288, 21906769, 265021, 264693, 35696423, 35695855, 264636, 56182323, 83373044, 87168518 |
| 3033 | 95308321 (6085, 6086) | Novel Protein sim. GBank gij5031573[ref]NP_005712.1pACTR - ARP3 (actin-related protein 3, yeast) homolog | Contains protein domain (PF00022) - Actin | struct | 35696286, 264259, 29331826, 35696032, 264508, 264905, 264908, 264907, 264908, 264909, 265008, 264591, 21906754, 265010, 265019, 264681, 264369, 264768, 21906764, 21906768, 35695917, 33657023, 264628, 35695855, 264632, 264635, 264639, 264482, 264563 |
| 3034 | 80415373 (6087, 6088) | | | UNCLASSIFIED | 264905, 264907, 264510, 264592, 265010, 264762, 264768, 264637, 264638, 264486 |
| 3035 | 91220692 (6089, 6070) | Novel Protein sim. GBank gij3738207[emb]CAA21262 - (AL031853) conserved ATP-GTP binding protein [Schizosaccharomyces pombe] | | UNCLASSIFIED | 264636 |
| 3036 | 91718323 (6071, 6072) | Novel Protein sim. GBank gij728837[sp]P39194/ALU7_HUMAN - IIII ALU SUBFAMILY SQ WARNING ENTRY IIII | | kinase | 264807, 33657402, 265021 |
| 3037 | 95307434 (6073, 6074) | Novel Protein sim. GBank gij4406590[gb]AAD20040 - (AF131766) Similar to Ena-VASP like protein [Homo sapiens] | | | 265017 |
| 3038 | 95421807 (6075, 6076) | Novel Protein sim. GBank gij5360093[gb]AAD42865.1[AF15509] NY-REN 18 antigen [Homo sapiens] | Contains protein domain (PF00627) - UBA domain | UNCLASSIFIED | 22278996, 22278997, 264259, 264905, 265007, 265009, 60433356, 21906754, 265018, 265019, 18108351, 264687, 21906765, 265020, 265021, 65274620, 27486262, 264636, 56182323, 18108385, 22279000 |
| 3039 | 87332257 (6077, 6078) | Novel Protein sim. GBank gij475712[emb]CAB42094.1 - (AJ238717) ZRP protein [Rattus norvegicus] | | UNCLASSIFIED | 35696286, 29331828, 264109, 264110, 264511, 265007, 21906754, 265011, 264681, 264683, 264687, 21906768, 264691, 18108370, 263972, 264629, 18108374, 263977, 35696423, 264584, 18108391, 264692, 264558, 18108382, 18108385, 264587 |
| 3040 | 90833317 (6079, 6080) | Novel Protein sim. GBank gij488427[emb]CAB43247.1 - (AL050037) hypothetical protein [Homo sapiens] | | | |
| 3041 | 88312357 (6081, 6082) | Novel Protein sim. GBank gij387607[emb]CAB04122.1 - (Z81505) similar to Zinc finger, C3HC4 type (RING finger); cDNA EST EMBL:D28025 comes from this gene; cDNA EST EMBL:D28024 comes from this gene; cDNA EST EMBL:D33210 comes from this gene; cDNA EST EMBL:D33441 comes from this ... | | UNCLASSIFIED | 56894075, 22278997, 22278998, 29331827, 33656970, 33109954, 21906754, 87168559, 264600, 264683, 21906765, 21906768, 22279002 |
| 3042 | 85749402 (6083, 6084) | Novel Protein sim. GBank gij790236 (U21156) - sarcolemmal associated protein-2 [Oryctolagus cuniculus] | | glycoprotein | 264636 |

| | | | | | |
|------|-----------------------|--|---|--------------|--|
| 3043 | 87773026 (6085, 6086) | Novel Protein sim. GBank gij854065[embjCAA58337] - (X83413) U88 [Human herpesvirus 6] | | UNCLASSIFIED | 3569286, 60424268, 35696052, 264508, 264905, 66712502, 56182435, 55811386, 52644296, 55811150, 35695917, 60170615, 33657109, 18108374, 264634, 60431850 |
| 3044 | 87646182 (6087, 6088) | Novel Protein sim. GBank gij4104922 (AF042276) - o251 homolog [Pseudomonas putida] | Contains protein domain (PF01209) - ubiE/COQ5 methyltransferase family | glycoprotein | 22278996, 22278998, 22278999, 29331824, 56182435, 264511, 265007, 60170831, 60432229, 60433356, 33109954, 18108351, 264288, 35695917, 18108388, 18108370, 60170394 |
| 3045 | 94127598 (6089, 6090) | Novel Protein sim. GBank gij4589680[dbjBAA76859.1] - (AB023232) KIAA1015 protein [Homo sapiens] | Contains protein domain (PF00096) - Zinc finger, C2H2 type | dna_rna_bind | 264488, 264259, 35696052, 264508, 264905, 264906, 264907, 264909, 264511, 265008, 264591, 264593, 33109954, 264604, 264764, 264683, 264288, 264766, 264768, 21906765, 21906768, 55811957, 35695917, 27486262, 18108370, 264628, 18108374, 35695855, 264630, 264632, 264635, 264563, 264564, 264566 |
| 3046 | 88088247 (6091, 6092) | | | UNCLASSIFIED | 22278999, 29331822, 29331824, 29331825, 29331826, 60432289, 29331827, 29331828, 264906, 52646317, 55811957, 60432113, 22279000, 22279002, 264482, 264564 |
| 3047 | 95089824 (6093, 6094) | | | UNCLASSIFIED | 264488, 22278996, 22278997, 22278999, 29331824, 29331825, 56182435, 264511, 265008, 265009, 265011, 265017, 264766, 21906768, 21906769, 35695917, 52644150, 33657349, 65274791, 35695855, 264555, 60432113, 22279000, 264566 |
| 3048 | 87629419 (6095, 6096) | Novel Protein sim. GBank gij4588034[gbjAAD25962.1]AF09287 - (AF092878) zinc RING finger protein SAG [Homo sapiens] | Contains protein domain (PF00097) - Zinc finger, C3HC4 type (RING finger) | UNCLASSIFIED | 264102, 29148784 |
| 3049 | 88229955 (6097, 6098) | Novel Protein sim. GBank gij5454158[pejNP_006286.1]pVARS - valyl-IRNA synthetase 1 | Contains protein domain (PF01406) - IRNA synthetases class I (C) | UNCLASSIFIED | 22278997, 29331826, 264807, 264758, 87168559, 265018, 264448, 21908766, 265020, 33657109, 35695855, 60432113, 22279000 |
| 3050 | 87643679 (6099, 6100) | Novel Protein sim. GBank gij4589642[dbjBAA76843.1] - (AB023216) KIAA0999 protein [Homo sapiens] | Contains protein domain (PF00069) - Eukaryotic protein kinase domain | kinase | 264259, 29331825, 264909, 265007, 264512, 265019, 264288, 21908766, 265020, 264693, 18108385, 56526486, 87168518, 22279002, 264566 |
| 3051 | 87750599 (6101, 6102) | | | | 22278997, 264595, 265018, 264288, 264693, 87168518 |
| 3052 | 57108030 (6103, 6104) | Novel Protein sim. GBank gij117528[spjP14755]CRYL_RABIT - LAMBDA-CRYSTALLIN | dehydrogenase | | 264534 |

| | | | | | |
|------|-----------------------|--|--|--------------|--|
| 3053 | 95350373 (6105, 6106) | Novel Protein sim. GBank gi 3947613 emb CAA19465.1 - (AL023828) cDNA EST EMBL:M89008 comes from this gene; cDNA EST yk282d3.5 comes from this gene [Caenorhabditis elegans] | | UNCLASSIFIED | 65274572, 56181686, 22278995, 35696286, 22278998, 264259, 60432289, 265008, 265009, 60433438, 21908754, 265010, 8718859, 264603, 265018, 265019, 264763, 264764, 264288, 21908765, 21908768, 21908768, 21908769, 35695917, 18108374, 35696423, 264638, 56182323, 22279000, 264563 |
| 3054 | 85843510 (6107, 6108) | Novel Protein sim. GBank gi 1076211 pir S50755 - hypothetical protein VSP-3 - Chlamydomonas reinhardtii | | UNCLASSIFIED | 35696286, 35696052, 28331830, 264908, 264909, 264512, 284910, 265017, 264604, 264766, 265020, 33657109, 264628, 35695855, 264636, 264564, 264566, 264486, 60424178, 65274572, 56182575, 35696286, 22278998, 22278999, 60432049, 264259, 60424269, 60432289, 35696052, 56182435, 265006, 265009, 60170831, 60432229, 60431735, 60433356, 264594, 60433438, 21908754, 55811386, 265011, 87168559, 265019, 18108351, 264683, 264288, 264369, 264689, 21908768, 55811957, 35695917, 60170815, 33657023, 65274620, 33657109, 35695763, 60431528, 18108374, 55810764, 55811576, 35696423, 65274791, 264636, 60431850, 18108381, 56182323, 60170394, 18108385, 60432113, 264564, 264565, 264566 |
| 3055 | 95350537 (6109, 6110) | Novel Protein sim. GBank gi 4680655 gb AAD27717.1 AF13294 - (AF132942) CGI-08 protein [Homo sapiens] | | transport | 264488, 264569, 18108394, 52646842, 22278997, 22278998, 22278999, 264259, 66714117, 29331826, 29331827, 35696052, 264508, 264509, 264905, 264906, 264907, 264908, 264909, 265006, 264512, 265007, 265008, 265009, 264910, 33657402, 55812038, 264596, 264758, 265010, 265011, 265017, 265019, 264760, 18108351, 264762, 264763, 264764, 264288, 264766, 264687, 18108357, 264768, 264768, 264689, 21908765, 21908766, 21908767, 21908768, 21908769, 35695917, 265020, 265021, 264691, 264693, 33657109, 18108370, 264628, 264629, 18108374, 55811576, 35696423, 35695855, 264630, 264631, 264632, 264634, 264635, 264636, 264637, 264638, 18108381, 83373044, 18108385, 22279000, 22279002, 264563, 264564, 264565, 264566, 264486, 264567 |
| 3056 | 91661636 (6111, 6112) | Novel Protein sim. GBank gi 728837 sp P39194 ALU7_HUMAN - IIII ALU SUBFAMILY SQ WARNING ENTRY IIII | | glycoprotein | |

| | | | | | |
|------|-----------------------|--|--|--------------|--|
| 3057 | 95412746 (6113, 6114) | Novel Protein sim. GBank gij3878119[embjCAA88860] - (Z49068) similar to GTP-binding protein; cDNA EST EMBL:M89111 comes from this gene; cDNA EST EMBL:D27709 comes from this gene; cDNA EST EMBL:D27708 comes from this gene; cDNA EST EMBL:D73788 comes from this gene; cDNA EST yk353... | | struct | 264508, 264905, 264907, 264908, 264909, 264510, 264512, 264910, 264592, 264594, 264767, 18108374, 264635, 264555, 264637, 264639, 264563, 264584, 264565, 264486 |
| 3058 | 79646226 (6115, 6116) | Novel Protein sim. GBank | | UNCLASSIFIED | 264693 |
| 3059 | 87629425 (6117, 6118) | gij4588034[gbjAAD25962.1]AF09287 - (AF092878) zinc RING finger protein SAG [Homo sapiens] | Contains protein domain (PF000097) - Zinc finger, C3HC4 type (RING finger) | UNCLASSIFIED | 22278995, 22278996, 22278997, 22278998, 22278999, 264480, 264259, 29331824, 29331825, 29331827, 35696052, 29331828, 285007, 60433438, 265017, 265018, 265019, 264681, 264448, 264288, 264768, 21906765, 21906766, 21906767, 21906769, 29148629, 29148784, 265022, 52644150, 18108370, 264636, 18108385, 264563, 264567 |
| 3060 | 79346691 (6119, 6120) | | | UNCLASSIFIED | 264567 |
| 3061 | 87740984 (6121, 6122) | | | UNCLASSIFIED | 264112, 52644296, 21906768, 33657023, 263974, 18108385 |
| 3062 | 87619465 (6123, 6124) | Novel Protein sim. GBank gij4454680[gbjAAD20963] - (AF070657) glutathione S-transferase subunit 13 homolog [Homo sapiens] | | transferase | 264908, 265008, 18108351, 264556 |
| 3063 | 80078023 (6125, 6126) | Novel Protein sim. GBank gij2246532 (U93872) - ORF 73, contains large complex repeat CR 73 [Kaposi's sarcoma-associated herpesvirus] | | UNCLASSIFIED | 18108359, 264558 |
| 3064 | 91241526 (6127, 6128) | Novel Protein sim. GBank gij4240315[gbjBAA74936.1] - (AB020720) KIAA0913 protein [Homo sapiens] | Contains protein domain (PF00403) - Heavy-metal-associated domain | UNCLASSIFIED | 52646365, 52646842, 65274572, 56182575, 58161686, 22278995, 22278996, 22278997, 22278998, 264259, 60432049, 29331824, 66714117, 264508, 264907, 264908, 56182435, 265009, 60432229, 60433438, 55812038, 52644296, 265018, 264682, 264288, 264686, 264768, 264687, 52644229, 264689, 21906768, 264691, 264692, 264693, 18108370, 18108377, 55811576, 264636, 56182323, 264558, 264639, 18108385, 22275000, 22279002 |
| 3065 | 91639201 (6129, 6130) | Novel Protein sim. GBank gij5656743[gbjAAD45960.1]AC00506 - (AC005067) Supported by Human EST H08032.1 (NID:g872854), mouse EST AA870042.1 (NID:g2965487), and gencan [Homo sapiens] | | UNCLASSIFIED | 22278996, 22278998, 264093, 264094, 264095, 29331824, 60424269, 66714117, 264100, 264907, 265007, 264591, 60432229, 264593, 265011, 265019, 18108351, 264766, 264767, 21906765, 21906768, 264693, 20281069, 22279000, 22279002, 264482, 264566, 264567 |

| | | | | | |
|------|-----------------------|---|--|--------------|--|
| 3066 | 91224437 (6131, 6132) | Novel Protein sim. GBank gj4884268[embjCAB43245.1] - (AL050028) hypothetical protein [Homo sapiens] | | UNCLASSIFIED | 18108397, 22278995, 56994075, 22278996, 264905, 68712502, 265008, 264512, 264910, 264758, 60174638, 264760, 18108351, 264764, 264683, 18108359, 264692, 18108364, 18108368, 18108370, 18108377, 18108379, 60170394, 264567 |
| 3067 | 85422551 (6133, 6134) | Novel Protein sim. GBank gj4689258[gbjAAD27832.1]AF12185 - (AF121859) sorting nexin 9 [Homo sapiens] | Contains protein domain (PF00787) - struct PX domain | | 264488, 264489, 35696286, 22278996, 56994075, 264259, 29331822, 29331825, 35696052, 29331828, 264508, 264905, 264509, 264908, 264907, 264908, 264909, 264112, 264510, 264511, 264512, 265008, 265009, 264910, 264591, 264592, 264593, 264594, 264757, 264595, 264596, 264758, 265010, 265011, 87168559, 264601, 264602, 264603, 264604, 264605, 265019, 264760, 264762, 264448, 264763, 264764, 264288, 264369, 264766, 264768, 264687, 264769, 264689, 21908765, 21908767, 21908768, 35695917, 265020, 265021, 264534, 52844150, 264691, 33657023, 264693, 264628, 60431528, 263977, 35695855, 264630, 264631, 264634, 264635, 264636, 264637, 264638, 264639, 83373044, 56526486, 87168518, 22279000, 22279002, 264563, 264483, 264564, 264565, 264566, 264567, 264486 |
| 3068 | 85360651 (6135, 6136) | Novel Protein sim. GBank gj3878119[embjCAA88860] - (Z49068) similar to GTP-binding protein; cDNA EST | Contains protein domain (PF01926) - struct GTPase of unknown function | | 22278996, 56994075, 22278998, 22278999, 264259, 264107, 264905, 29331830, 52844045, 264110, 60170831, 264592, 264594, 33657402, 21908764, 33109954, 87168474, 87168559, 265017, 264448, 264764, 264683, 264766, 52644229, 21908765, 21908766, 21908768, 21908769, 60170615, 33657023, 18108370, 18108376, 264634, 264557, 60170394, 56182323, 18108385, 87168518, 22279000, 264482 |
| 3069 | 85412753 (6137, 6138) | EMBL: M89111 comes from this gene; cDNA EST EMBL: D27709 comes from this gene; cDNA EST EMBL: D27708 comes from this gene; cDNA EST EMBL: D73788 comes from this gene; cDNA EST yk353... | | | |

| | | | | | |
|------|-----------------------|--|--|------------------|--|
| 3070 | 94319173 (6139, 6140) | Novel Protein sim. GBank gj 3877788 emb CAB05527 - (Z83110) cDNA EST yk472b5.3 comes from this gene; cDNA EST yk474a7.3 comes from this gene; cDNA EST yk472b5.5 comes from this gene; cDNA EST yk468c10.3 comes from this gene; cDNA EST yk468c10.5 comes from this gene; cDNA EST EM... | | synthase | 264488, 22278994, 22278995, 22278996, 56994075, 22278997, 22278999, 264259, 29331822, 29147620, 29331824, 66714117, 29331826, 29146498, 29146499, 66712502, 29331830, 52644045, 56182435, 264511, 265007, 264512, 264910, 60170831, 264592, 264758, 33109954, 21906754, 87168474, 265019, 18108351, 264448, 264683, 264288, 52644229, 264689, 21906765, 21906766, 21906767, 21906769, 35695917, 265020, 265021, 60170815, 52644150, 264691, 33657023, 27486261, 27486264, 264628, 33657023, 27486261, 27486264, 264628, 18108370, 18108377, 55811576, 35695955, 264634, 264635, 18108381, 60170394, 56182323, 264558, 83373044, 18108385, 18108387, 56526486, 264404, 264563, 264566 |
| 3071 | 94325573 (6141, 6142) | Novel Protein sim. GBank gj 4502425 ref NP_001709.1 pBMP6 - bone morphogenetic protein 6 precursor | Contains protein domain (PF00085) - Iqf Thioredoxin | | 264488, 65274572, 18108398, 22278996, 35696286, 22278997, 22278998, 22278999, 264259, 29331822, 66714117, 29331826, 35696052, 29331828, 29146498, 29146499, 264907, 264908, 29331830, 264909, 52644045, 56182435, 265006, 265007, 264512, 265008, 265009, 60170831, 60432228, 264592, 60433356, 33657402, 60433438, 33109954, 52644296, 87168474, 265010, 265017, 264681, 264288, 264685, 264766, 264687, 264769, 264689, 21906765, 21906766, 21906767, 21906768, 21906769, 35695917, 265020, 265021, 265022, 60170815, 52644150, 264690, 264691, 264692, 33657023, 264693, 33657109, 263971, 18108374, 18108377, 35696423, 55811576, 65274791, 35695855, 264630, 264635, 264636, 264557, 60170394, 83373044, 60432113, 22279000, 22279002, 264563, 264564, 264565, 264566, 264567, 56182575, 29331822, 29331824, 29331825, 29146498, 264908, 52644045, 56182435, 265009, 60433438, 55812038, 18108351, 264683, 264369, 52644229, 52644150, 33657023, 264693, 33657109, 18108374, 55811576, 65274791, 264555, 56182323, 60432113, 264564 |
| 3072 | 95115892 (6143, 6144) | Novel Protein sim. GBank gj 1263289 (U47856) - fibroin-4 [Araneus diadematus] | | transcriptfactor | |

| | | | | | |
|------|-----------------------|--|---|--------------|--|
| 3073 | 85147248 (6145, 6146) | Novel Protein sim. GBank gi134840 sp P22528 CORB_HUMAN - CORNIFIN B (SMALL PROLINE-RICH PROTEIN IB) (SPR-IB) (14.8 KD PANCORNUIN) | | UNCLASSIFIED | 264769 |
| 3074 | 88089351 (6147, 6148) | Novel Protein sim. GBank gi3419847 (AC004982) - similar to yeast hypothetical protein ybk4; similar to P38164 (PID:g586461) [Homo sapiens] | | UNCLASSIFIED | 264488, 265019, 264448, 264288, 21906767, 264893, 18108388, 18108370, 18108374, 264567 |
| 3075 | 88085752 (6148, 6150) | Novel Protein sim. GBank gi4557349 ref NP_000458.1 pBARD - BRCA1 associated RING domain 1 | Contains protein domain (PF00023) - Ank repeat | homeobox | 264509, 264907, 264689, 264893, 56526486 |
| 3076 | 87819219 (6151, 6152) | | | UNCLASSIFIED | 18108398, 29331822, 29331827, 60432229, 265017, 264691, 264693 |
| 3077 | 88734277 (6153, 6154) | Novel Protein sim. GBank gi3023956 sp Q00808 HET1_PODAN - VEGETATIBLE INCOMPATIBILITY PROTEIN HET-E-1 | Contains protein domain (PF00400) - WD domain, G-beta repeat | kinase | 65274572, 35696052, 264511, 60170831, 87188474, 264389, 35695917, 33657182, 27486264, 33657349, 35695763, 35695855, 264639 |
| 3078 | 88089355 (6155, 6156) | Novel Protein sim. GBank gi3900850 (AC004994) - similar to KIAA0800; similar to d1026456 (PID:g3043724) [Homo sapiens] | | | 22279002 |
| 3079 | 87821893 (6157, 6158) | Novel Protein sim. GBank gi3875410 emb CAB02876 - (Z81052) Similarity to Yeast ABC1P protein (SW:ABC1_YEAST); cDNA EST yk229g8.3 comes from this gene; cDNA EST yk229g8.5 comes from this gene [Caenorhabditis elegans] | | transport | 29331824, 29331826, 264758, 55811386, 265017, 55811150, 52644229, 21906768, 265020, 265021, 264693, 18108376, 264631, 52644332, 22279002 |
| 3080 | 93298274 (6159, 6160) | Novel Protein sim. GBank gi5257221 gb AAD41265.1 - (AF117887) protein arginine methyltransferase [Mus musculus] | | interferon | 264488, 52644507, 22278996, 22278998, 264490, 264259, 29331824, 6671417, 29331825, 29331826, 29331827, 29331828, 29146499, 264508, 264905, 264828, 52644045, 56182435, 265006, 264591, 264596, 21908754, 60174639, 265010, 264882, 264448, 264763, 264683, 264764, 264288, 264685, 264769, 264688, 264689, 21906765, 21906767, 21906769, 55811957, 35695917, 265020, 60170615, 52644150, 264692, 33657023, 264693, 65274620, 33657109, 27486261, 35695763, 264628, 18108370, 65274791, 264558, 56182323, 60170394, 264482, 264565, 264484 |
| 3081 | 88094864 (6161, 6162) | Novel Protein sim. GBank gi128831 sp P39188 ALU1_HUMAN - IIII ALU SUBFAMILY J WARNING ENTRY IIII | | UNCLASSIFIED | 18108398, 264509, 264905, 264906, 264907, 264908, 264909, 264510, 264511, 265009, 264910, 264595, 264758, 265011, 265018, 264760, 264761, 264763, 264764, 18108354, 264685, 264766, 264628, 264629, 264630, 264631, 264632, 264634, 264635, 264555, 264638, 18108382, 18108385, 264563, 264565, 264566 |
| 3082 | 80310121 (6163, 6164) | | | | 264764, 55811957, 264555, 264564 |

| | | | | | |
|------|-----------------------|--|---|--------------|--|
| 3083 | 88095756 (6165, 6166) | Novel Protein sim. GBank gi 868241 (U29488) - C56C10.3 gene product [Caenorhabditis elegans] | | UNCLASSIFIED | 264488, 264259, 29331824, 264106, 265008, 264591, 264592, 21906754, 264288, 264767, 21906768, 21906769, 29148784, 264691, 264632, 22279000 |
| 3084 | 87448568 (6167, 6168) | Novel Protein sim. GBank gi 476774 (p1rjA37475 - probable structural component p38 - borna disease virus | | | 22278995, 60432289, 35696032, 264905, 264906, 264907, 264908, 264909, 265006, 265007, 264910, 264593, 264595, 264758, 264369, 264288, 264766, 35695917, 265020, 18108374, 35696423, 264631, 264556, 264555, 264586, 264587, 264488 |
| 3085 | 87795781 (6169, 6170) | Novel Protein sim. GBank gi 2565057 (U80741) - CAGH44 [Homo sapiens] | | UNCLASSIFIED | 265011, 264681 |
| 3086 | 87769942 (6171, 6172) | Novel Protein sim. GBank gi 3894189 (AC005662) - hypothetical protein [Arabidopsis thaliana] | | UNCLASSIFIED | 22278998, 264092, 264259, 29331822, 29331825, 264108, 264112, 18108351, 264687, 263987, 263974, 55810764, 263981, 18108385, 264487 |
| 3087 | 87462888 (6173, 6174) | | | | 52646365, 56994075, 22278997, 22278998, 29331824, 29331825, 35696052, 60433438, 33109954, 21906754, 52646317, 265017, 264682, 264369, 264684, 21906767, 21906768, 265020, 264691, 33657023, 33657109, 52645129, 33657182, 27486262, 35695855, 87168518 |
| 3088 | 91224441 (6175, 6176) | Novel Protein sim. GBank gi 3355304 (AF001549) - Unknown gene product [Homo sapiens] | | UNCLASSIFIED | 264591 |
| 3089 | 95361242 (6177, 6178) | Novel Protein sim. GBank gi 4689146 (p1AAD27782.1) (AF077049) lambda-crystallin [Homo sapiens] | Contains protein domain (PF00725) - dehydrogenase 3-hydroxyacyl-CoA dehydrogenase | | 18108397, 65274572, 56182575, 56181686, 56994075, 35696286, 22278997, 22278998, 264259, 29331824, 29331825, 29331826, 29331828, 264907, 29331830, 264909, 56182435, 264510, 265007, 60170831, 60432228, 21906754, 55811386, 265017, 265018, 265019, 264760, 55811150, 264288, 264766, 56181562, 21906765, 21906766, 21906767, 21906768, 265021, 60170615, 27486262, 18108370, 60431528, 35696423, 264558, 264559, 60432113, 264486 |

| | | | | | |
|------|-----------------------|---|--|--------------|---|
| 3090 | 95342371 (6179, 6180) | Novel Protein sim. GBank gij1354050 (U47024) - MEM3 [Mus musculus] | | UNCLASSIFIED | 60424178, 52645156, 65274572, 56182575, 56181686, 22278995, 35696286, 56994075, 22278996, 22278998, 22278999, 264259, 29331822, 56182181, 29331824, 29331825, 29331826, 29331827, 29331828, 35696052, 33656970, 264908, 264908, 52644045, 264828, 265006, 265007, 265008, 60170831, 60432229, 60433356, 33657402, 55812038, 264758, 21906754, 33109954, 52646317, 55811388, 52644298, 87168474, 265011, 87168559, 265017, 265018, 265019, 55811150, 18108351, 264681, 264448, 264288, 264369, 18108357, 264768, 52644229, 56181562, 21906764, 21906765, 21906766, 21906767, 21906768, 21906769, 35695917, 265020, 265022, 60170615, 264650, 52644150, 264691, 33657023, 18108365, 65274620, 33657109, 18108368, 33657182, 27486261, 27486265, 35695763, 18108374, 18108376, 55810764, 35696423, 55811576, 65274791, 35695855, 264557, 56182323, 83373044, 18108387, 18108388, 87168518, 22278000, 22279002, 264563, 264482 |
| 3091 | 95317424 (6181, 6182) | Novel Protein sim. GBank gij3873932[emb]CAB01859] - (Z79598) Similarity to Bovine aspartyl beta hydroxylase (TR:G162694); cDNA EST EMBL:D27916 comes from this gene; cDNA EST EMBL:D27915 comes from this gene; cDNA EST EMBL:D64881 comes from this gene; cDNA EST EMBL:D68139 comes f... | | UNCLASSIFIED | 35696286, 29331822, 35696052, 264508, 264509, 264905, 264906, 264908, 264909, 264510, 264758, 265010, 265011, 264683, 264685, 264766, 264768, 264769, 264693, 264628, 35696423, 35695855, 264632, 264635, 264639, 264482, 264563, 264486 |

| | | | | | |
|------|-----------------------|---|--|------------------|---|
| 3092 | 95314592 (6183, 6184) | Novel Protein sim. GBank gi11710756 sp P15880 RS2_HUMAN - 40S RIBOSOMAL PROTEIN S2 (S4) (LLREP3 PROTEIN) | Contains protein domain (PF00333) - Ribosomal protein S5 | - ribosomal prot | 264488, 60424179, 18108396, 22278995, 56994075, 22278998, 35698286, 22278997, 22278998, 60432049, 264259, 29331822, 29331824, 29331825, 29331826, 29331827, 35686052, 29331828, 29146498, 29146499, 284508, 284509, 284905, 284906, 264907, 29331830, 264908, 284909, 264113, 264510, 264511, 265006, 264512, 265007, 265008, 264910, 265009, 60170831, 264591, 264592, 60431735, 264593, 264594, 60433438, 284595, 264758, 21908754, 265010, 265011, 264601, 264602, 265017, 264603, 264604, 265018, 264605, 265019, 264760, 264762, 264681, 18108351, 264763, 264682, 264448, 264764, 264683, 264288, 264369, 264765, 264766, 264686, 264767, 264687, 264768, 264769, 264688, 21906764, 264689, 21906765, 21906766, 21906767, 21906768, 21906769, 29148629, 29148784, 35895917, 265020, 265021, 264534, 60170615, 284690, 264691, 264692, 65274620, 33657109, 27486262, 264628, 264629, 18108374, 263978, 18108377, 35696423, 264630, 264631, 264632, 264634, 264635, 264555, 264636, 264637, 264556, 264638, 264557, 264558, 264639, 60170394, 18108385, 264259, 29331824, 35696052, 264905, 285008, 60432229, 60431735, 264684, 264369, 264288, 264766, 21906767, 35696423, 83373044, 18108385 |
| 3093 | 94318457 (6185, 6186) | Novel Protein sim. GBank gi15002587 emb CAB44347.1 - (Y17454) LSR1 protein [Homo sapiens] | UNCLASSIFIED | | |
| 3094 | 94316875 (6187, 6188) | Novel Protein sim. GBank gi1400734 sp P31044 PBP_RAT - PHOSPHATIDYLETHANOLAMINE-BINDING PROTEIN (23 KD MORPHINE-BINDING PROTEIN) (P23K) | Contains protein domain (PF01161) - Phosphatidylethanolamine-binding protein | collagen | 18108398, 264259, 60432289, 29331827, 264511, 264763, 264288, 264767, 265022, 264691, 264693, 65274791, 56182323, 264564, 264565 |

| | | | | | |
|------|-----------------------|--|---|--------------|---|
| 3095 | 94848182 (6189, 6180) | Novel Protein sim. GBank gij487759jgb/AAD31421.1 AF12444 - (AF124440) MAGE tumor antigen D1 [Homo sapiens] | Contains protein domain (PF01454) - MAGE family | UNCLASSIFIED | 18108397, 56182575, 22278995, 35696286, 56994075, 22278997, 22278999, 264259, 60432049, 66714117, 29331825, 60432289, 35696052, 33656970, 29146499, 264508, 264905, 264509, 29331830, 264909, 264510, 264511, 264512, 265007, 265008, 265009, 60170831, 264758, 21906754, 85658542, 265010, 265011, 87168559, 265017, 265018, 265019, 264760, 264681, 264682, 264683, 264764, 264369, 264288, 264686, 264768, 264769, 264689, 21906765, 21906766, 21906767, 55811957, 35695917, 265020, 265021, 265022, 52644150, 264691, 264692, 33657023, 264693, 263972, 18108376, 55811576, 35696423, 264952, 60170394, 264639, 83373044, 18108385, 18108387, 65274727, 87168518, 60432113, 264482, 264563, 264584, 264566, 264487, 18108391, 22278995, 22278996, 22278997, 22278999, 29331824, 29331825, 29331826, 29331827, 33656970, 264905, 264908, 265008, 264910, 33657402, 265011, 265017, 265018, 264369, 21906766, 21906767, 21906768, 35695917, 265020, 60170615, 264691, 264692, 264693, 27486261, 27486262, 18108370, 60431528, 264634, 264636, 264639, 22279000, 264566, 264488, 29331822, 29331825, 60432289, 29331826, 35696052, 29331828, 29331830, 264594, 55812038, 33109954, 33657084, 87168474, 87168559, 52644229, 21906765, 21906767, 18108376, 35696423, 52644332, 264638, 60432113, 22279002 |
| 3096 | 87756128 (6191, 6182) | Novel Protein sim. GBank gij3882221 dbj BAA34470.1 - (AB018293) KIAA0750 protein [Homo sapiens] | Contains protein domain (PF00307) - sitrucl Calponin homology (CH) domain | UNCLASSIFIED | |
| 3097 | 88264895 (6193, 6194) | Novel Protein sim. GBank gij4468288 emb CAB37981 - (AL022395) dJ273N12.1 (PUTATIVE protein based on EST matches) [Homo sapiens] | Contains protein domain (PF00846) - F-box domain. | UNCLASSIFIED | |
| 3098 | 80258024 (6195, 6186) | | | | 264634, 264637, 264565 |
| 3099 | 91243325 (6197, 6198) | Novel Protein sim. GBank gij303603 dbj BAA02145.1 - (D12621) cytochrome P-450L TBV [Homo sapiens] | | cyto450 | 264488, 35696286, 29331822, 29331824, 29331825, 29331827, 265007, 265008, 265010, 265011, 265018, 265019, 18108357, 21906766, 265020, 265022, 55811576, 56182323, 22279002, 264563 |
| 3100 | 87602421 (6199, 6200) | Novel Protein sim. GBank gij1083764 pri B48013 - proline- rich proteoglycan 2 precursor, parotid - rat | | UNCLASSIFIED | 29331825, 60432289, 35696052, 264910, 60432229, 264592, 264288, 264693, 263967, 264635 |
| 3101 | 79602134 (6201, 6202) | | | UNCLASSIFIED | 264908, 264693, 264628, 264630, 264632 |

| | | | | | |
|------|-----------------------|--|--|--------------|--|
| 3102 | 91220892 (6203, 6204) | Novel Protein sim. GBank gi 5305708 gb AAD41781.1 AF12853 - (AF128536) cytoplasmic phosphoprotein PACSIN2 [Homo sapiens] | Contains protein domain (PF00018) - SH3 domain | - struct | 35696286, 22278996, 22278999, 29331827, 35696052, 264909, 264512, 265008, 601170831, 60433356, 33109954, 18108351, 264684, 264689, 21908767, 60170615, 264692, 33657023, 264638, 22279000, 264482, 264564 |
| 3103 | 80938004 (6205, 6206) | Novel Protein sim. GBank gi 464584 sp P35292 RB17_MOUSE - RAS-RELATED PROTEIN RAB-17 | | UNCLASSIFIED | 35695917, 264565 |
| 3104 | 87340633 (6207, 6208) | Novel Protein sim. GBank gi 5032207 ref NP_005696.1 pTSSC - tumor-suppressing STF cDNA 6 | | UNCLASSIFIED | 264259, 264684, 264532, 33657182, 264558 |
| 3105 | 94148603 (6209, 6210) | | | | 22278997, 264259, 29331824, 35696052, 29331828, 264508, 264509, 264905, 264906, 264907, 264908, 264511, 264910, 264591, 264594, 264758, 264760, 264681, 264762, 264764, 264288, 264766, 264768, 264687, 264769, 21906766, 21906768, 35695917, 33657023, 264692, 264693, 264628, 264629, 35695855, 264630, 264631, 264632, 264634, 264635, 264637, 264638, 264639, 83373044, 264404, 22279002, 264563, 264565, 264566, 264486, 264567 |
| 3106 | 95361416 (6211, 6212) | Novel Protein sim. GBank gi 1838574 (U97190) - B0025.2 gene product [Caenorhabditis elegans] | | | 22278996, 22278997, 22278998, 22278999, 264092, 264093, 264094, 29331822, 264906, 264907, 264908, 52644045, 56182435, 264112, 265008, 265009, 55812038, 265017, 265018, 264683, 264686, 264687, 264768, 52644229, 21906765, 21906768, 21906769, 55811957, 265020, 265022, 264690, 52644150, 264692, 264693, 18108370, 18108377, 55811576, 56182323, 18108385, 18108388, 22279000, 264563 |
| 3107 | 95343272 (6213, 6214) | Novel Protein sim. GBank gi 3341441 emb CAA76851 - (Y17794) winged-helix transcription factor [Gallus gallus] | | | 22278995, 22278996, 35696286, 22278997, 22278999, 264091, 264093, 264259, 29331822, 29331825, 29331826, 60432289, 29331827, 29331828, 33656970, 264105, 264512, 265009, 60433356, 60433438, 265011, 265017, 265018, 21906765, 21906766, 21906767, 21906769, 265021, 264691, 33657109, 27486261, 27486265, 18108370, 263972, 18108374, 55811576, 18108385, 56526488, 264482, 264487, 56182435, 264288, 264690, 264564 |
| 3108 | 87340635 (6215, 6216) | Novel Protein sim. GBank gi 5032207 ref NP_005696.1 pTSSC - tumor-suppressing STF cDNA 6 | | UNCLASSIFIED | |

| | | | | | |
|------|-----------------------|---|--|------------------|--|
| 3109 | 94318461 (6217, 6218) | Novel Protein sim. GBank gii5002587[emb]CAB44347.1] - (Y17454) LSFRI protein [Homo sapiens] | Contains protein domain (PF00096) - Zinc finger, C2H2 type | struct | 264480, 264908, 265007, 264910, 264593, 264683, 264684, 264687, 21906767, 21906768, 264693, 18108370, 264629, 18108374, 264632, 264636, 22279000 |
| 3110 | 95090716 (6218, 6220) | Novel Protein sim. GBank gii1076211[pir]S50755 - hypothetical protein VSP-3 - Chlamydomonas reinhardtii | UNCLASSIFIED | | 264488, 65274572, 22278995, 22278997, 60432049, 264259, 29331822, 29331824, 29331825, 60432288, 29331826, 29331827, 29331828, 264908, 264510, 265008, 265007, 265008, 265009, 60432229, 33857402, 60433356, 265011, 87168559, 264600, 265017, 265018, 265019, 18108351, 264288, 264369, 21906766, 21906767, 21906768, 265020, 60170615, 264693, 65274620, 18108370, 264638, 18108384, 22279000, 264563, 18108390 |
| 3111 | 87754512 (6221, 6222) | Novel Protein sim. GBank gii3282231 (U75454) - C2H2 type zinc finger protein [Homo sapiens] | Contains protein domain (PF00096) - Zinc finger, C2H2 type | transcriptfactor | 264488, 18108398, 66712502, 265017, 265018, 265019, 264448, 21906767, 265020, 33657023, 18108365, 18108368, 35686423, 52644332, 18108385, 18108388 |
| 3112 | 88043639 (6223, 6224) | Novel Protein sim. GBank gii3900848 (AC005023) - match to EST AA361117 (NID: g2013436) [Homo sapiens] | Contains protein domain (PF00046) - Homeobox domain | homeobox | |
| 3113 | 88207098 (6225, 6226) | Novel Protein sim. GBank gii2459910 (AF005856) - anon2A5 [Drosophila yakuba] | Homeobox domain | lm7 | 18108397, 22278899, 264259, 29331824, 35686052, 264907, 264757, 60433438, 87168559, 264763, 264448, 18108354, 264288, 21906767, 21906769, 35695917, 264690, 264691, 264692, 264693, 18108365, 18108381, 18108384, 18108385, 18108388, 87168518, 22279000, 22279002 |
| 3114 | 79843167 (6227, 6228) | Novel Protein sim. GBank gii4968270[gb]AAB52261.2] - (U97002) similar to acyl-CoA dehydrogenases and epoxide hydrolases; Pfam domain PF00441 (Acyl-CoA_dh). Score=57.4, E-value=1.7e-16, N=2; contains similarity to Pfam domain PF00702 (Hydrolase), Score=57.4, E-value=1e-13, N=1 [C.... | Contains protein domain (PF00702) - haloacid dehalogenase-like hydrolase | hydrolase | |
| 3115 | 94117896 (6229, 6230) | Novel Protein sim. GBank gii5032225[re]NP_005676.1[wp]BSC - Williams-Beuren syndrome chromosome region 11 | | transcriptfactor | 60424179, 56182575, 264259, 29331824, 60424269, 29331826, 66712502, 264510, 265007, 60431735, 60433356, 55812038, 55811386, 265019, 264288, 264689, 21906769, 264691, 33657023, 264693, 60431528, 263974, 60431850, 56182323, 264559, 22279000, 22279002 |
| 3116 | 79842855 (6231, 6232) | | | UNCLASSIFIED | 264905, 264758, 21906764, 264690 |
| 3117 | 87771288 (6233, 6234) | | | UNCLASSIFIED | 264510, 265011, 18108351, 264288, 264689, 264691, 18108368, 18108372, 263981, 264558, 264564 |

| | | | | | |
|------|-----------------------|--|--|---------------|--|
| 3118 | 94655848 (6235, 6236) | Novel Protein sim. GBank gi 3880563 emb CAB01444.1 - (Z78016) predicted using GeneFinder; similar to serine/threonine kinase; cDNA EST YK353d10.5 comes from this gene [Caenorhabditis elegans] | Contains protein domain (PF00008) - EGF-like domain | Igf | 52645156, 52646842, 65274572, 56182575, 22278995, 56994075, 22278996, 35696286, 22278997, 22278998, 22278999, 264259, 29331822, 29331824, 66714117, 29331826, 29331827, 35696052, 29331828, 264905, 264908, 29331830, 52644045, 56182435, 264510, 264511, 265007, 265008, 265009, 264757, 52646317, 21906754, 33657084, 52644298, 87168474, 87168559, 265017, 265018, 264605, 265019, 264762, 264448, 264682, 264684, 264288, 264766, 56181562, 21906765, 21906766, 21906769, 265020, 265022, 264690, 52644150, 264691, 33657023, 264693, 33657109, 33657349, 264628, 18108370, 60431528, 18108374, 35698423, 65274791, 60170394, 83373044, 87168518, 22279000, 22279002, 264486, 265006, 264288 |
| 3119 | 85728796 (6237, 6238) | | Contains protein domain (PF00328) - Histidine acid phosphatase | | |
| 3120 | 87344040 (6238, 6240) | Novel Protein sim. GBank gi 5019819 gb AAD37863.1 AF14315 - (AF143152) putative NADH oxidoreductase complex I subunit [Caenorhabditis elegans] | | UNCLASSIFIED | 264488, 264509, 264510, 264511, 264512, 264288, 264486 |
| 3121 | 94110735 (6241, 6242) | Novel Protein sim. GBank gi 4501877 ref NP_001088.1 pACRQ - acrosin | | | 52644507, 52645156, 52646365, 52646842, 22278994, 56994075, 22278996, 22278999, 264259, 29331824, 29331827, 35696052, 52644045, 265008, 52646317, 87168474, 87168559, 21906765, 52644150, 33657023, 18108374, 264637 |
| 3122 | 11814528 (6243, 6244) | | | UNCLASSIFIED | 264638 |
| 3123 | 8803003 (6245, 6246) | Novel Protein sim. GBank gi 2439517 (AC002563) - putative RHO/RAC effector protein; 85% similarity to P49205 (PIDg1345860) [Homo sapiens] | Contains protein domain (PF00780) - CNH domain | kinase | 18108392, 29331822, 29331824, 29331825, 264905, 265007, 55812038, 265019, 18108351, 264682, 264288, 264766, 21906764, 21906765, 21906768, 21906769, 55811957, 18108365, 18108366, 27486265, 18108374, 18108381, 18108384, 22279000, 22279002, 264482 |
| 3124 | 87766899 (6247, 6248) | | | UNCLASSIFIED | 264905 |
| 3125 | 91216607 (6249, 6250) | Novel Protein sim. GBank gi 4980828 gb AAD35412.1 AE00171 - (AE001714) oxidoreductase, short chain dehydrogenase/reductase family [Thermotoga maritima] | Contains protein domain (PF00106) - short chain dehydrogenase | dehydrogenase | 56181686, 264259, 66714117, 60432289, 29331826, 29331827, 264907, 264908, 264828, 265009, 60433356, 33657402, 60433438, 264758, 18108351, 264288, 29148627, 29148629, 33657023, 33657109, 18108382, 56526486 |

| | | | | | | |
|------|-----------------------|--|---|--|--------------|--|
| 3126 | 85337205 (6251, 6252) | | | | UNCLASSIFIED | 22278999, 264490, 264259, 60432049, 29331822, 60432289, 29146498, 52644045, 56182435, 265009, 60433438, 265010, 87168559, 265017, 265018, 55811150, 264763, 264683, 264368, 264685, 28148629, 33657023, 264693, 33657109, 18108374, 55811578, 18108385, 60432113, 22279002, 35696286, 22278999, 28331828, 264808, 60433438, 87168559, 264604, 21906765, 21906769, 33657023, 33657349, 264629, 18108374, 18108377, 22279000, 22279002 |
| 3127 | 91639233 (6253, 6254) | Novel Protein sim. GBank gi 2828280 emb CAA16694.1 - (AL021887) putative protein [Arabidopsis thaliana] | | | | |
| 3128 | 87674330 (6255, 6256) | Novel Protein sim. GBank gi 3885828 (AF090133) - lin-7-A [Rattus norvegicus] | Contains protein domain (PF00595) - PDZ domain (Also known as DHR or GLGF). | | misc_channel | 22278998, 264259, 52644045, 265008, 21906754, 265017, 265018, 21906768, 18108376, 18108387, 22279000, 22279002 |
| 3129 | 87755412 (6257, 6258) | Novel Protein sim. GBank gi 3135273 (AC003058) - hypothetical protein [Arabidopsis thaliana] | Contains protein domain (PF00400) - WD domain, G-beta repeat | | kinase | 56182575, 264259, 29331825, 29331828, 52644045, 56182435, 60433356, 264600, 264682, 264763, 264784, 264369, 264288, 264686, 55811957, 264692, 33657023, 33657109, 60432113, 264564, 264566, 264636 |
| 3130 | 14983860 (6259, 6260) | Novel Protein sim. GBank gi 3329485 (AF064553) - NSD1 protein [Mus musculus] | | | | |
| 3131 | 85351469 (6261, 6262) | Novel Protein sim. GBank gi 1848277 (U86136) - telomerase-associated protein TP-1 [Homo sapiens] | Contains protein domain (PF00400) - WD domain, G-beta repeat | | UNCLASSIFIED | 56182575, 264259, 29331824, 264907, 56182435, 264594, 60433438, 55812038, 33109954, 21906754, 33657084, 87168474, 264448, 264766, 21906769, 55811957, 265020, 265021, 265022, 60170615, 33657023, 33657109, 33657182, 27486261, 33657349, 65274791, 60170394, 56182323, 83373044, 87168518, 264564 |

| | | | | |
|------|-----------------------|--|---|---|
| 3132 | 95415459 (6263, 6264) | Novel Protein sim. GBank gi 4680647 gb AAD27713.1 AF13293 - (AF132938) CGI-03 protein [Homo sapiens] | Contains protein domain (PF00789) - ubiquitin UBX domain | 52644507, 52646842, 52646365, 5274572, 56182575, 22278994, 22278995, 35986286, 56994075, 22278998, 22278997, 22278998, 22278999, 60432049, 52645080, 29331822, 29331824, 29331825, 60432289, 29331826, 29331827, 29331828, 35696052, 264508, 52644045, 56182435, 264910, 60170831, 60432229, 60433356, 33657402, 55812038, 52646317, 21908754, 52644296, 85658542, 87168559, 265017, 265018, 265019, 264448, 264288, 264389, 52644229, 21906765, 21906766, 21906767, 21906768, 21906769, 35695917, 265020, 265021, 60170615, 52644150, 264692, 33657023, 52845129, 33657109, 33657182, 27486261, 27486262, 27486265, 33657349, 35695763, 18108374, 18108376, 55811576, 35695855, 18108365, 18108387, 56526486, 87168518, 60432113, 22279002 |
| 3133 | 87379414 (6265, 6266) | Novel Protein sim. GBank gi 4507613 ref NP_003738.1 pTNKS - TANKYRASE | polymerase | 22278994, 22278998, 264905, 265006, 265007, 87168559, 264760, 21906767, 18108374, 22279000, 22279002, 264563 264595, 264369, 264685, 264628, 264566 |
| 3134 | 94649816 (6267, 6268) | Novel Protein sim. GBank gi 1729827 sp P54633 TALA_DICDI - FILOPODIN (ITALIN HOMOLOG) | | 22278998, 264095, 29331826, 33657402, 18108348, 263974 |
| 3135 | 86389356 (6269, 6270) | Novel Protein sim. GBank gi 3093478 (AF012927) - fibrinogen-binding protein [Streptococcus equi] | struct | 22278998, 264259, 264828, 265006, 265008, 60433438, 265019, 264764, 264288, 264769, 264699, 265020, 27486262, 263972, 65274791, 264557, 264558 |
| 3136 | 94845839 (6271, 6272) | Novel Protein sim. GBank gi 627101 pir IS44092 - probable carrier protein c2 - Caenorhabditis elegans | Contains protein domain (PF00153) - Mitochondrial carrier proteins | 22278995, 22278996, 22278997, 22278998, 22278999, 264259, 29331822, 29331825, 29331826, 29331827, 29331828, 264510, 265008, 21906754, 87168474, 265011, 87168559, 265017, 265018, 265019, 18108351, 264682, 264769, 21906765, 21906766, 21906767, 21906769, 55811857, 35695917, 265020, 265021, 52844150, 18108370, 18108374, 22279000, 22279002, 264482, 264486 |
| 3137 | 88257947 (6273, 6274) | Novel Protein sim. GBank gi 3342730 (AC005331) - R31341_1 [Homo sapiens] | UNCLASSIFIED | |

| | | | | | |
|------|-----------------------|--|--------------|--|--|
| 3138 | 94130186 (6275, 6276) | Novel Protein sim. GBank gii4406759[gbjAAD20070] - (AC006638) hypothetical protein [Arabidopsis thaliana] | | | 264569, 264488, 264907, 264511, 264593, 33109854, 87168559, 264681, 264684, 264685, 264686, 264687, 264768, 264688, 264689, 264691, 264692, 264693, 33657109, 264631, 264634, 264635, 264636, 264637, 60170394, 83373044, 18108385, 18108388, 60432113, 22279000, 22279002 |
| 3139 | 87325503 (6277, 6278) | Novel Protein sim. GBank gii228938[prfj]1814452C - Hyp-rich glycoprotein [Zea diploperennis] | UNCLASSIFIED | | 22278997, 22278998, 22278999, 264905, 265018, 265019, 21906765, 265020, 264636, 264557 |
| 3140 | 91222692 (6278, 6280) | Novel Protein sim. GBank gii932[embjCAA37773] - (X53744) 68kDa subunit of signal recognition particle [Canis familiaris] | struct | | 22278995, 56994075, 35696286, 264908, 264909, 60433358, 21906754, 52644296, 87168474, 87168559, 264683, 264288, 264685, 264686, 265022, 264693, 27486262, 35695855, 264630, 264555, 264566 |
| 3141 | 87323564 (6281, 6282) | Novel Protein sim. GBank gii3213227 (AF035209) - putative v-SNARE Vti1a [Mus musculus] | UNCLASSIFIED | | 56182575, 35696286, 29331828, 264909, 265009, 265018, 18108351, 264369, 21906766, 29148627, 265020, 264628, 264629, 264631, 18108385 |
| 3142 | 95419028 (6283, 6284) | Novel Protein sim. GBank gii2498197[spjQ95245][C561_PIG - CYTOCHROME B561 (CYTOCHROME B-561)] | cytochrome | | 52645156, 52646365, 22278995, 35696286, 22278998, 22278999, 60432049, 264259, 29331822, 29331824, 29331827, 29146489, 56182435, 265007, 60170831, 60432229, 33657402, 264595, 60433438, 264758, 21906754, 264288, 264768, 264687, 52644229, 21906765, 21906767, 21808768, 60170615, 52644150, 65274620, 33657109, 35695763, 18108370, 18108376, 65274791, 35695855, 264631, 264557, 87168518, 60432113, 22279000 |
| 3143 | 95351475 (6285, 6286) | Novel Protein sim. GBank gii5420387[embjCAB46679.1] - (AJ243459) proteophosphoglycan [Leishmania major] | UNCLASSIFIED | | 264488, 56182575, 22278998, 22278999, 29331822, 29331824, 60432289, 35696052, 29331828, 264508, 264905, 264906, 264907, 264908, 264909, 52644045, 56182435, 264511, 264512, 265008, 264910, 60432229, 33657402, 60433356, 60433438, 55812038, 265011, 265019, 264760, 264763, 264448, 264764, 264684, 264288, 264685, 264686, 264768, 264689, 21906765, 21906766, 21906767, 21906769, 35695917, 264690, 33657023, 264693, 263987, 33657109, 264628, 264629, 18108374, 263976, 55811576, 35695855, 264630, 264631, 264632, 264634, 264635, 264636, 264637, 264558, 87168518, 60432113, 22279000, 22279002, 264563, 264566, 264486 |

| | | | | | |
|------|-----------------------|--|--|--------------|--|
| 3144 | 95336329 (6287, 6288) | Novel Protein sim. GBank gll488446[emb]CAB43322.1] - (AL050225) hypothetical protein [Homo sapiens] | | | 264488, 18108396, 22278996, 35696286, 22278997, 22278999, 29331828, 29331827, 35696052, 29331828, 264106, 265006, 265007, 265009, 33657402, 85658542, 265011, 18108351, 264448, 264369, 21908765, 21908766, 21908767, 265020, 265021, 52644150, 27486261, 18108370, 18108374, 35696423, 56182323, 83373044, 22279000, 22279002, 264567 |
| 3145 | 86611657 (6289, 6290) | Novel Protein sim. GBank gll387970[emb]CAB03330] - (Z81118) Similarity to Human endosomal protein P162 (TR:Q15075); cDNA EST EMBL:Z14487 comes from this gene; cDNA EST EMBL:Z14556 comes from this gene; cDNA EST EMBL:D27011 comes from this gene; cDNA EST EMBL:D27015 comes from this gene | UNCLASSIFIED | | 18108397, 29331824, 29146499, 20281100, 265006, 55812038, 265010, 21908766, 29148627, 21908768, 29148784, 264692, 33657023, 33657109, 35695763, 263981, 56182323, 87168518 |
| 3146 | 87756314 (6291, 6292) | Novel Protein sim. GBank gll2135746[pir]S69890 - mitogen inducible gene mig-2 - human | Contains protein domain (PF00169) - PH domain | struct | 264259, 29331826, 29331828, 29331830, 264510, 264511, 265007, 265009, 264600, 265017, 18108351, 264448, 264369, 21908766, 265021, 264692, 33657109, 18108374, 35696423, 35695855, 60432113, 264564 |
| 3147 | 94848512 (6293, 6294) | Novel Protein sim. GBank gll387427[emb]CAB07315.1] - (Z92825) predicted using GeneFinder; cDNA EST yk315e12.3 comes from this gene; cDNA EST yk315e12.5 comes from this gene [Caenorhabditis elegans] | Contains protein domain (PF00702) - haloacid dehalogenase-like hydrolase | UNCLASSIFIED | 56181686, 35696286, 60432049, 264259, 56182181, 29331825, 60432289, 35696052, 56182435, 265008, 264910, 60431735, 60433356, 60433438, 265010, 264448, 264288, 265022, 33657023, 33657109, 60431528, 65274791, 264631, 56182323, 264404, 22279002 |
| 3148 | 95382169 (6295, 6296) | Novel Protein sim. GBank gll5225322[gb]AAD40851.1[AF08310] - (AF083108) sirutin type 3 [Homo sapiens] | UNCLASSIFIED | | 35696286, 35696052, 264511, 85658542, 87168474, 264764, 35696423, 264555, 264556, 264557, 264558, 83373044, 56526486, 60432113 |
| 3149 | 95308548 (6297, 6298) | Novel Protein sim. GBank gll4200446 (AF102777) - FYVE finger-containing phosphoinositide kinase [Mus musculus] | Contains protein domain (PF01363) - FYVE zinc finger | eph | 29331822, 35696052, 264109, 29148628, 18108381 |
| 3150 | 87655472 (6299, 6300) | Novel Protein sim. GBank gll3378454[emb]CAA76893] - (Y17850) ganglioside-induced differentiation associated protein 1 [Mus musculus] | Contains protein domain (PF00043) - Glutathione S-transferases | transferase | 264259, 29331822, 29331824, 29331825, 29331827, 52646317, 264686, 35695855, 56182323, 264639 |
| 3151 | 87772355 (6301, 6302) | Novel Protein sim. GBank gll172591 (M63577) - SFP1 [Saccharomyces cerevisiae] | Contains protein domain (PF00096) - Zinc finger, C2H2 type | oncogene | 29331822, 265008 |
| 3152 | 85698108 (6303, 6304) | | UNCLASSIFIED | | 21906754, 87168559, 264605, 21906768, 52644150, 27486264, 35696423, 22279000 |

| | | | | | |
|------|-----------------------|---|--|-------------------|--|
| 3153 | 95317299 (6305, 6308) | Novel Protein sim. GBank gi 4895041 gb AAD32705.1 AF143957) coronin-3 [Mus musculus] | Contains protein domain (PF00400) - struct WD domain, G-beta repeat | - struct | 284488, 52646385, 35698286, 22278996, 22278997, 22278999, 60432049, 264259, 29331826, 60432289, 33658970, 264508, 264908, 33657402, 264595, 60433438, 87168474, 87168559, 264601, 265019, 264448, 264682, 264764, 264288, 264369, 264768, 21906765, 21906768, 21906767, 21906768, 21906769, 29148784, 265021, 265022, 60170615, 52644150, 264690, 264691, 33657023, 65274620, 33657109, 18108370, 35695855, 264638, 60170394, 87168518, 60432113, 22279000, 22279002, 22278998, 264259, 29331824, 66712502, 265008, 265010, 265017, 18108354, 264691, 33657023, 264693, 20281149, 18108374, 29331828, 264509, 264905, 264908, 264510, 264511, 264512, 33657402, 264681, 264683, 33657023, 18108370, 264634, 264639, 18108385, 264563, 264488 |
| 3154 | 87718573 (6307, 6308) | Novel Protein sim. GBank gi 4880661 gb AAD27720.1 AF132945) CGI-11 protein [Homo sapiens] | ATPase associated | ATPase associated | 56182575, 22278996, 22278997, 22278998, 22278999, 60432049, 264259, 29331822, 29331824, 66714117, 29331825, 29331826, 29331827, 35696052, 52644045, 265007, 265009, 60170831, 60432229, 60433356, 21906754, 33109954, 87168474, 265010, 265017, 265018, 265019, 18108351, 264448, 264288, 264689, 21906768, 21906768, 21906769, 35695917, 265020, 265022, 264692, 18108370, 35698423, 56182323, 22279002 |
| 3155 | 87722394 (6309, 6310) | Novel Protein sim. GBank gi 728637 sp P39184 ALU7_HUMAN - IIII ALU SUBFAMILY SQ WARNING ENTRY IIII | UNCLASSIFIED | UNCLASSIFIED | 21808396, 264259, 29331826, 35696052, 29146498, 87168559, 265017, 264448, 264288, 264691, 18108366, 52645129, 35698423, 52644332 |
| 3156 | 87737449 (6311, 6312) | Novel Protein sim. GBank gi 5830078 gb AAD45821.1 AC006017) N-acetylglucosaminyltransferase, similar to Q10473 (PID:g1709559) [Homo sapiens] | Contains protein domain (PF00652) - Similarity to lectin domain of ricin beta-chain, 3 copies. | transferase | 264488, 263974 |
| 3157 | 88259577 (6313, 6314) | | | | |
| 3158 | 80034118 (6315, 6316) | Novel Protein sim. GBank gi 5306064 gb AAD41895.1 AF156778) ASB-3 protein [Homo sapiens] | Contains protein domain (PF00023) - Ank repeat | kinase | |
| 3159 | 84124114 (6317, 6318) | Novel Protein sim. GBank gi 5531272 emb CAB50897.1 - (AJ243800) WSC4 homologue [Kluyveromyces fragilis] | UNCLASSIFIED | UNCLASSIFIED | 56182575, 22278998, 29331824, 264108, 60433356, 264758, 265011, 87168559, 264448, 18108354, 264768, 21906768, 265020, 264691, 264692, 33657109, 18108374, 35698423, 264555, 60170394, 22279000 |
| 3160 | 80221088 (6319, 6320) | Novel Protein sim. GBank gi 3930525 AF064447) - sex-determination protein homolog Fem1a [Mus musculus] | Contains protein domain (PF00023) - Ank repeat | - struct | 18108351, 264555, 264556, 264557, 264558, 264559 |

| | | | | | |
|------|----------------------|--|--|--|---|
| 3161 | 8807411 (6321, 6322) | | | | 264488, 2227895, 2227897, 2227898, 264259, 2933182, 6043289, 2933182, 5264045, 265017, 265018, 264448, 264288, 21908784, 21908787, 265020, 18108374, 264636, 264568 |
|------|----------------------|--|--|--|---|

Table 2

| Tissue ID | Tissue Name | Tissue Information | Disease Association |
|-----------|-------------------------------|--------------------------------|---|
| 20281069 | 192xN | Protein-protein Interactions | Any |
| 20281071 | 192xN | Protein-protein Interactions | Any |
| 20281149 | 192xN | Protein-protein Interactions | Any |
| 20281152 | 192xN | Protein-protein Interactions | Any |
| 264111 | 276xN | Protein-protein Interactions | Any |
| 264112 | 276xN | Protein-protein Interactions | Any |
| 263966 | 384xN | Protein-protein Interactions | Any |
| 263967 | 384xN | Protein-protein Interactions | Any |
| 264110 | 552xN | Protein-protein Interactions | Any |
| 18108379 | SPH 52.1 (Adrenal Gland) | Adrenal Gland/Suprarenal gland | Adrenoleukodystrophy , Congenital Adrenal Hyperplasia, |
| 18108381 | SPH 52.2 (Fetal Lung) | Fetal Lung | Cystic Fibrosis, infection, lung cancer |
| 18108383 | SPH 52.3 (B's Lymphoma- Raji) | Lymphoma derived from B cells | Blood cancers, hematopoiesis, leukemia |
| 18108368 | SPH 52.5 (Salivary Gland) | Salivary Gland | Dry mouth, infection |
| 18108384 | SPH 52.6 (Brain- Thalmus) | Thalamus | Brain cancer, head injury, obesity, neurological disorders, neuropsychiatric disorders |
| 18108394 | SPH 53.1 (Adrenal Gland) | Adrenal Gland/Suprarenal gland | Adrenoleukodystrophy , Congenital Adrenal Hyperplasia, |
| 18108355 | SPH 53.2 (Fetal Lung) | Fetal Lung | Cystic Fibrosis, infection, lung cancer |
| 18108359 | SPH 53.3 (B's Lymphoma- Raji) | Lymphoma derived from B cells | Blood cancers, hematopoiesis, leukemia |
| 18108361 | SPH 53.4 (Mammary Gland) | Mammary Gland | Lactation disorders, breast cancer |
| 18108362 | SPH 53.5 (Salivary Gland) | Salivary Gland | Dry mouth, infection |
| 18108366 | SPH 53.6 (Brain- Thalmus) | Thalamus | Brain cancer, head injury, obesity, neurological disorders, neuropsychiatric disorders |
| 18108354 | SPH 54.1 (Adrenal Gland) | Adrenal Gland/Suprarenal gland | Adrenoleukodystrophy , Congenital Adrenal Hyperplasia, |
| 18108392 | SPH 54.2 (Fetal Lung) | Fetal Lung | Cystic Fibrosis, infection, lung cancer |
| 18108348 | SPH 54.3 (B's Lymphoma- Raji) | Lymphoma derived from B cells | Blood cancers, hematopoiesis, leukemia |
| 18108382 | SPH 54.4 (Mammary Gland) | Mammary Gland | Lactation disorders, breast cancer |
| 18108395 | SPH 54.5 (Salivary Gland) | Salivary Gland | Dry mouth, infection |
| 18108365 | SPH 54.6 (Brain- Thalmus) | Thalamus | Brain cancer, head injury, obesity, neurological disorders, neuropsychiatric disorders |
| 18108397 | SPH 55.1 (Adrenal Gland) | Adrenal Gland/Suprarenal gland | Adrenoleukodystrophy , Congenital Adrenal Hyperplasia, |
| 18108398 | SPH 55.2 (Fetal Lung) | Fetal Lung | Cystic Fibrosis, infection, lung cancer |
| 18108364 | SPH 55.3 (B's Lymphoma- Raji) | Lymphoma derived from B cells | |
| 18108388 | SPH 55.4 (Mammary Gland) | Mammary Gland | Lactation disorders, breast cancer |
| 18108358 | SPH 55.5 (Salivary Gland) | Salivary Gland | Dry mouth, infection |
| 20281099 | SPH 56.2 (MG63) | | |
| 20281100 | SPH 56.3 (U1SMC) | | |
| 264404 | SPH.1 (Brain) | Whole Brain | Von Hippel-Lindau (VHL) syndrome , Alzheimer's disease, Stroke, Tuberous sclerosis, hypercalcaemia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxia-telangiectasia, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuroprotection |

| | | | |
|--------|--|--------------------------|--|
| 264510 | SPH.10 (Pancreas) | Pancreas | Pancreatitis, diabetes, pancreatic cancer |
| 264511 | SPH.11 (Placenta) | Placenta | Infertility, birth defects |
| 264512 | SPH.12 (Thyroid) | Thyroid | Hyperparathyroidism, Hypoparathyroidism |
| 264555 | SPH.13 (Bone Marrow) | Bone Marrow | Hemophilia, hypercoagulation, Idiopathic thrombocytopenic purpura, autoimmune disease, allergies, immunodeficiencies, transplantation, Graft versus host, |
| 264556 | SPH.14 (Bone Marrow) | Bone Marrow | Hemophilia, hypercoagulation, Idiopathic thrombocytopenic purpura, autoimmune disease, allergies, immunodeficiencies, transplantation, Graft versus host, |
| 264557 | SPH.15 (Bone Marrow) | Bone Marrow | Hemophilia, hypercoagulation, Idiopathic thrombocytopenic purpura, autoimmune disease, allergies, immunodeficiencies, transplantation, Graft versus host, |
| 264558 | SPH.16 (Bone Marrow) | Bone Marrow | Hemophilia, hypercoagulation, Idiopathic thrombocytopenic purpura, autoimmune disease, allergies, immunodeficiencies, transplantation, Graft versus host, |
| 264559 | SPH.17 (Bone Marrow) | Bone Marrow | Hemophilia, hypercoagulation, Idiopathic thrombocytopenic purpura, autoimmune disease, allergies, immunodeficiencies, transplantation, Graft versus host, |
| 264569 | SPH.19 (One Fetal tissue and two cell lines) | Mixed | |
| 264687 | SPH.19.1 (fetal thymus - CRL7046) | Fetal Thymus | Hemophilia, hypercoagulation, Idiopathic thrombocytopenic purpura, immunodeficiencies |
| 264688 | SPH.19.2 (hematopoietic stem cells - CRL2043) | Hematopoietic stem cells | Leukemia, osteoporosis, post-chemotherapeutic stem cell repopulation |
| 264689 | SPH.19.3 (osteogenic sarcoma cell lines - HTB96) | Osteogenic Sarcoma | Sarcomas, osteoporosis, osteopetrosis |
| 264690 | SPH.19.4 (Fetal Liver) | Fetal liver | Von Hippel-Lindau (VHL) syndrome, Cirrhosis, Transplantation |
| 264691 | SPH.19.5 (Heart) | Heart | Cardiomyopathy, Atherosclerosis, Hypertension, Congenital heart defects, Aortic stenosis, Atrial septal defect (ASD), Atrioventricular (A-V) canal defect, Ductus arteriosus, Pulmonary stenosis, Subaortic stenosis, Ventricular septal defect (VSD), valve diseases, Tuberous sclerosis, Scleroderma, Obesity, Transplantation |
| 264692 | SPH.19.6 (Spleen) | Spleen | Hemophilia, Hypercoagulation, Idiopathic thrombocytopenic purpura, Immunodeficiencies, Graft versus host |
| 264693 | SPH.19.7 (Pituitary) | Pituitary | Von Hippel-Lindau (VHL) syndrome, Alzheimer's disease, Stroke, Tuberous sclerosis, hypercalcaemia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxia-telangiectasia, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuroprotection |
| 264482 | SPH.2 (Brain) | Brain | Von Hippel-Lindau (VHL) syndrome, Alzheimer's disease, Stroke, Tuberous sclerosis, hypercalcaemia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxia-telangiectasia, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuroprotection |

| | | | |
|--------|------------------------|--------------|---|
| 264600 | SPH.21 (Fetal Brain) | Fetal brain | Von Hippel-Lindau (VHL) syndrome , Alzheimer's disease, Stroke, Tuberous sclerosis, hypercalcaemia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxia-telangiectasia, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuroprotection |
| 264601 | SPH.22 (Bone Marrow) | Bone Marrow | Hemophilia, hypercoagulation, Idiopathic thrombocytopenic purpura, autoimmune disease, allergies, immunodeficiencies, transplantation, Graft versus host, |
| 264602 | SPH.23 (Thyroid) | Thyroid | Hyperthyroidism and Hypothyroidism |
| 264603 | SPH.24 (Pancreas) | Pancreas | Pancreatitis, diabetes, pancreatic cancer |
| 264604 | SPH.25 (Lymph Node) | Lymph Node | Lymphedema , Allergies |
| 264605 | SPH.26 (Placenta) | Placenta | Infertility, birth defects |
| 264634 | SPH.28 (Heart) | Heart | Cardiomyopathy, Atherosclerosis, Hypertension, Congenital heart defects, Aortic stenosis , Atrial septal defect (ASD), Atrioventricular (A-V) canal defect, Ductus arteriosus , Pulmonary stenosis , Subaortic stenosis, Ventricular septal defect (VSD), valve diseases, Tuberous sclerosis, Scleroderma, Obesity, Transplantation |
| 264635 | SPH.29 (Fetal Kidney) | Fetal Kidney | Diabetes, Autoimmune disease, Renal artery stenosis, Interstitial nephritis, Glomerulonephritis, Polycystic kidney disease, Systemic lupus erythematosus, Renal tubular acidosis, IgA nephropathy, Hypercalcaemia, Lesch-Nyhan syndrome |
| 264483 | SPH.3 (Bone Marrow) | Bone marrow | Hemophilia, hypercoagulation, Idiopathic thrombocytopenic purpura, autoimmune disease, allergies, immunodeficiencies, transplantation, Graft versus host, |
| 264636 | SPH.30 (Lymph Node) | Lymph Node | Lymphedema , Allergies |
| 264637 | SPH.31 (Pancreas) | Pancreas | Pancreatitis, diabetes, pancreatic cancer |
| 264638 | SPH.32 (Thyroid) | Thyroid | Hyperthyroidism and Hypothyroidism |
| 264639 | SPH.33 (Fetal Brain) | Fetal brain | Von Hippel-Lindau (VHL) syndrome , Alzheimer's disease, Stroke, Tuberous sclerosis, hypercalcaemia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxia-telangiectasia, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuroprotection |
| 264484 | SPH.4 (Bone Marrow) | Bone Marrow | Hemophilia, hypercoagulation, Idiopathic thrombocytopenic purpura, autoimmune disease, allergies, immunodeficiencies, transplantation, Graft versus host, |
| 264758 | SPH.44.1 (Kidney) | Kidney | Diabetes, Autoimmune disease, Renal artery stenosis, Interstitial nephritis, Glomerulonephritis, Polycystic kidney disease, Systemic lupus erythematosus, Renal tubular acidosis, IgA nephropathy, Hypercalcaemia, Lesch-Nyhan syndrome |
| 264760 | SPH.44.2 (Fetal Liver) | Fetal Liver | Von Hippel-Lindau (VHL) syndrome, Cirrhosis, Transplantation |

| | | | |
|----------|-------------------------------------|--------------------------------|--|
| 264762 | SPH.44.3 (Heart) | Heart | Cardiomyopathy, Atherosclerosis, Hypertension, Congenital heart defects, Aortic stenosis, Atrial septal defect (ASD), Atrioventricular (A-V) canal defect, Ductus arteriosus, Pulmonary stenosis, Subaortic stenosis, Ventricular septal defect (VSD), valve diseases, Tuberous sclerosis, Scleroderma, Obesity, Transplantation |
| 264764 | SPH.44.4 (Prostate) | Prostate | Prostate Cancer |
| 264766 | SPH.44.5 (Spleen) | Spleen | Hemophilia, Hypercoagulation, Idiopathic thrombocytopenic purpura, Immunodeficiencies, Graft versus host |
| 264768 | SPH.44.6 (pituitary) | Pituitary | Von Hippel-Lindau (VHL) syndrome, Alzheimer's disease, Stroke, Tuberous sclerosis, hypercalcaemia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxia-telangiectasia, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuroprotection |
| 264769 | SPH.44.7 (Uterus) | Uterus | Infertility, birth defects |
| 264905 | SPH.48.1 (Burkitt's Lymphoma- Raji) | Burkitt's Lymphoma | Lymphoma, blood cancers |
| 264906 | SPH.48.2 (Thalamus- Brain) | Thalamus | Von Hippel-Lindau (VHL) syndrome, Alzheimer's disease, Stroke, Tuberous sclerosis, hypercalcaemia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxia-telangiectasia, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuroprotection |
| 264907 | SPH.48.3 (Adrenal Gland) | Adrenal Gland/Suprarenal gland | Adrenoleukodystrophy, Congenital Adrenal Hyperplasia, |
| 264908 | SPH.48.4 (Fetal Lung) | Fetal Lung | Cystic Fibrosis, infection, lung cancer |
| 264909 | SPH.48.5 (Salivary Gland) | Salivary Gland | Dry mouth, infection |
| 264910 | SPH.48.6 (Mammary Gland) | Mammary Gland | Lactation disorders, breast cancer |
| 265006 | SPH.50.1 (B's lymphoma) | Burkitt's Lymphoma | Lymphoma, blood cancers |
| 265007 | SPH.50.2 (thalamus) | Thalamus | Von Hippel-Lindau (VHL) syndrome, Alzheimer's disease, Stroke, Tuberous sclerosis, hypercalcaemia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxia-telangiectasia, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuroprotection |
| 265008 | SPH.50.3 (adrenal gland) | Adrenal Gland/Suprarenal gland | Adrenoleukodystrophy, Congenital Adrenal Hyperplasia, |
| 265009 | SPH.50.4 (fetal lung) | Fetal Lung | Cystic Fibrosis, infection, lung cancer |
| 265010 | SPH.50.5 (salivary gland) | Salivary Gland | Dry mouth, infection |
| 265011 | SPH.50.6 (mammary gland) | Mammary Gland | Lactation disorders, breast cancer |
| 18108385 | SPH.51.1 (MCF-7) | Breast Cancer | Breast Cancer |
| 18108370 | SPH.51.2 (CCRF-CEM) | Cancer Cell line | Cancer |
| 18108374 | SPH.51.3 (K-562) | Cancer Cell line | Cancer |
| 18108351 | SPH.51.4 (OVCAR-3) | Ovarian cancer | Ovarian cancer |
| 18108372 | SPH.51.5 (HL-60) | Cancer Cell line | Cancer |
| 264486 | SPH.6 (Bone Marrow) | Bone Marrow | Hemophilia, hypercoagulation, Idiopathic thrombocytopenic purpura, autoimmune disease, allergies, immunodeficiencies, transplantation, Graft versus host, |

| | | | |
|----------|----------------------|-------------|---|
| 264508 | SPH.8 (Fetal Brain) | | Von Hippel-Lindau (VHL) syndrome , Alzheimer's disease, Stroke, Tuberous sclerosis, hypercalcaemia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxia-telangiectasia, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuroprotection |
| 264509 | SPH.9 (Lymph Node) | Lymph Node | Lymphedema , Allergies |
| 20798451 | SRH 56.3 (U1SMC) | | |
| 264487 | SRH.1 (Brain) | Brain | Von Hippel-Lindau (VHL) syndrome , Alzheimer's disease, Stroke, Tuberous sclerosis, hypercalcaemia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxia-telangiectasia, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuroprotection |
| 264534 | SRH.11 (Bone marrow) | Bone Marrow | Hemophilia, hypercoagulation, Idiopathic thrombocytopenic purpura, autoimmune disease, allergies, immunodeficiencies, transplantation, Graft versus host, |
| 264535 | SRH.12 (Bone marrow) | Bone Marrow | Hemophilia, hypercoagulation, Idiopathic thrombocytopenic purpura, autoimmune disease, allergies, immunodeficiencies, transplantation, Graft versus host, |
| 264563 | SRH.19 (Fetal Brain) | Fetal brain | Von Hippel-Lindau (VHL) syndrome , Alzheimer's disease, Stroke, Tuberous sclerosis, hypercalcaemia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxia-telangiectasia, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuroprotection |
| 264488 | SRH.2 (Bone Marrow) | Bone Marrow | Hemophilia, hypercoagulation, Idiopathic thrombocytopenic purpura, autoimmune disease, allergies, immunodeficiencies, transplantation, Graft versus host, |
| 264564 | SRH.20 (Lymph Node) | Lymph Node | Lymphedema , Allergies |
| 264565 | SRH.21 (Pancreas) | Pancreas | Pancreatitis, diabetes, pancreatic cancer |
| 264566 | SRH.22 (Placenta) | Placenta | Infertility, birth defects |
| 264567 | SRH.23 (Thyroid) | Thyroid | Hyperthyroidism and Hypothyroidism |
| 264591 | SRH.25 (Fetal Brain) | Fetal brain | Von Hippel-Lindau (VHL) syndrome , Alzheimer's disease, Stroke, Tuberous sclerosis, hypercalcaemia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxia-telangiectasia, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuroprotection |
| 264592 | SRH.26 (Bone Marrow) | Bone Marrow | Hemophilia, hypercoagulation, Idiopathic thrombocytopenic purpura, autoimmune disease, allergies, immunodeficiencies, transplantation, Graft versus host, |
| 264593 | SRH.27 (thyroid) | Thyroid | Hyperthyroidism and Hypothyroidism |
| 264594 | SRH.28 (Pancreas) | Pancreas | Pancreatitis, diabetes, pancreatic cancer |
| 264595 | SRH.29 (Lymph Node) | Lymph Node | Lymphedema , Allergies |
| 264489 | SRH.3 (Bone Marrow) | Bone Marrow | Hemophilia, hypercoagulation, Idiopathic thrombocytopenic purpura, autoimmune disease, allergies, immunodeficiencies, transplantation, Graft versus host, |

| | | | |
|--------|--|--------------------------|--|
| 264596 | SRH.30 (Placenta) | Placenta | Infertility, birth defects |
| 264628 | SRH.33 (fetal Kidney) | Fetal kidney | Diabetes, Autoimmune disease, Renal artery stenosis, Interstitial nephritis, Glomerulonephritis, Polycystic kidney disease, Systemic lupus erythematosus, Renal tubular acidosis, IgA nephropathy, Hypercalcaemia, Lesch-Nyhan syndrome |
| 264629 | SRH.34 (lymph Node) | Lymph Node | Lymphedema , Allergies |
| 264630 | SRH.35 (Pancreas) | Pancreas | Pancreatitis, diabetes, pancreatic cancer |
| 264631 | SRH.36 (thyroid) | Thyroid | Hyperthyroidism and Hypothyroidism |
| 264632 | SRH.37 (Fetal Brain) | Fetal Brain | Von Hippel-Lindau (VHL) syndrome , Alzheimer's disease, Stroke, Tuberous sclerosis, hypercalcaemia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxia-telangiectasia, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuroprotection |
| 264490 | SRH.4 (Bone Marrow) | Bone Marrow | Hemophilia, hypercoagulation, Idiopathic thrombocytopenic purpura, autoimmune disease, allergies, immunodeficiencies, transplantation, Graft versus host, |
| 264681 | SRH.43.1 (fetal thymus - CRL7046) | Fetal Thymus | Hemophilia, hypercoagulation, Idiopathic thrombocytopenic purpura, immunodeficiencies |
| 264682 | SRH.43.2 (hematopoietic stem cells - CRL2043) | Hematopoietic stem cells | Leukemia, osteoporosis, post-chemotherapeutic stem cell repopulation |
| 264683 | SRH.43.3 (osteogenic sarcoma cell lines - HTB96) | Osteogenic Sarcoma | Sarcomas, osteoporosis, osteopetrosis |
| 264684 | SRH.43.4 (Fetal Liver) | Fetal Liver | Von Hippel-Lindau (VHL) syndrome, Cirrhosis, Transplantation |
| 264685 | SRH.43.6 (Spleen) | Spleen | Hemophilia, Hypercoagulation, Idiopathic thrombocytopenic purpura , Immunodeficiencies, Graft versus host |
| 264686 | SRH.43.7 (pituitary) | Pituitary | Von Hippel-Lindau (VHL) syndrome , Alzheimer's disease, Stroke, Tuberous sclerosis, hypercalcaemia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxia-telangiectasia, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuroprotection |
| 264757 | SRH.44.1 (Kidney) | Kidney | Diabetes, Autoimmune disease, Renal artery stenosis, Interstitial nephritis, Glomerulonephritis, Polycystic kidney disease, Systemic lupus erythematosus, Renal tubular acidosis, IgA nephropathy, Hypercalcaemia, Lesch-Nyhan syndrome |
| 264759 | SRH.44.2 (Fetal Liver) | Fetal Liver | Von Hippel-Lindau (VHL) syndrome, Cirrhosis, Transplantation |
| 264761 | SRH.44.3 (Heart) | Heart | Cardiomyopathy, Atherosclerosis, Hypertension, Congenital heart defects, Aortic stenosis, Atrial septal defect (ASD), Atrioventricular (A-V) canal defect, Ductus arteriosus , Pulmonary stenosis , Subaortic stenosis, Ventricular septal defect (VSD), valve diseases, Tuberous sclerosis, Scleroderma, Obesity, Transplantation |
| 264763 | SRH.44.4 (Prostate) | Prostate | Prostate Cancer |
| 264765 | SRH.44.5 (Spleen) | Spleen | Hemophilia, Hypercoagulation, Idiopathic thrombocytopenic purpura , Immunodeficiencies, Graft versus host |

| | | | |
|----------|------------------------------|--------------------------------|--|
| 264767 | SRH.44.6 (Pituitary) | Pituitary | Von Hippel-Lindau (VHL) syndrome , Alzheimer's disease, Stroke, Tuberous sclerosis, hypercalcaemia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxia-telangiectasia, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuroprotection |
| 264828 | SRH.46.1 (Lymph Node) | Lymph Node | Lymphedema , Allergies |
| 264887 | SRH.47.5 (Fetal Liver) | Fetal Liver | Von Hippel-Lindau (VHL) syndrome, Cirrhosis, Transplantation |
| 18108377 | SRH.50.1 (B's lymphoma) | Burkitt's Lymphoma | Lymphoma, blood cancers |
| 18108380 | SRH.50.2 (thalamus) | Thalamus | Von Hippel-Lindau (VHL) syndrome , Alzheimer's disease, Stroke, Tuberous sclerosis, hypercalcaemia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxia-telangiectasia, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuroprotection |
| 18108396 | SRH.50.3 (adrenal gland) | Adrenal Gland/Suprarenal gland | Adrenoleukodystrophy , Congenital Adrenal Hyperplasia. |
| 18108391 | SRH.50.4 (fetal lung) | Fetal Lung | Airway diseases, infection |
| 18108357 | SRH.50.5 (salivary gland) | Salivary Gland | Dry mouth, infection |
| 18108390 | SRH.50.6 (mammary gland) | Mammary Gland | Lactation disorders, breast cancer |
| 264532 | SRH.9 (Bone Marrow) | Bone Marrow | Hemophilia, hypercoagulation, Idiopathic thrombocytopenic purpura, autoimmune disease, allergies, immunodeficiencies, transplantation, Graft versus host, |
| 263974 | 736xN | | |
| 263976 | 736xN | | |
| 263981 | 736xN | | |
| 20281166 | 96xN | | |
| 20281169 | 96xN | | |
| 20281171 | 96xN | | |
| 263994 | cDNA-ORF Selection | | |
| 264080 | Mx96 | | |
| 21906754 | NQH 6.1 (HH729) | | |
| 22278996 | NQH 6.10 (PrEC) | Endothelial cells | heart disease, cancer |
| 22278997 | NQH 6.11 (CAEC) | Endothelial cells | heart disease, cancer |
| 22278998 | NQH 6.12 (CSC) | Cancer Cell line | Cancer |
| 22278999 | NQH 6.13 (NHNPC) | Cancer Cell line | Cancer |
| 22279000 | NQH 6.14 (NHMC-RM) | Cancer Cell line | Cancer |
| 22279002 | NQH 6.15 (Hypothalamus) | Hypothalamus | Von Hippel-Lindau (VHL) syndrome , Alzheimer's disease, Stroke, Tuberous sclerosis, hypercalcaemia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxia-telangiectasia, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuroprotection, Obesity |
| 21906764 | NQH 6.2 (In Dated Platelets) | Platelets | Clotting diseases, stroke |
| 21906765 | NQH 6.3 (HuVec) | Endothelial cells | heart disease, cancer |
| 87168474 | NQH 6.3 (Sized-HUVEC) | Endothelial cells | heart disease, cancer |
| 21906766 | NQH 6.4 (UuMVEC- myo) | Cancer Cell line | Cancer |
| 21906767 | NQH 6.5 (NHEM-neo) | Cancer Cell line | Cancer |
| 21906768 | NQH 6.6 (NHEK) | Cancer Cell line | Cancer |
| 21906769 | NQH 6.7 (ByCAEC) | Endothelial cells | heart disease, cancer |
| 22278994 | NQH 6.8 (NHA) | Cancer Cell line | Cancer |

| | | | |
|----------|--------------------------------------|-------------------|---|
| 22278995 | NQH 6.9 (PrSC) | Cancer Cell line | Cancer |
| 27486261 | NQH 7.1 (Jurkat E6-untreated) | Cancer Cell line | Cancer |
| 27486262 | NQH 7.2 (TF1-untreated) | Cancer Cell line | Cancer |
| 27486264 | NQH 7.3 (U87-untreated) | Cancer Cell line | Cancer |
| 27486265 | NQH 7.4 (THP1-untreated) | Cancer Cell line | Cancer |
| 29331822 | NQH 8.1 (Brain- amygdala) | | Von Hippel-Lindau (VHL) syndrome , Alzheimer's disease, Stroke, Tuberous sclerosis, hypercalcaemia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxia-telangiectasia, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuroprotection |
| 29331824 | NQH 8.2 (Brain-hippocampus) | | Von Hippel-Lindau (VHL) syndrome , Alzheimer's disease, Stroke, Tuberous sclerosis, hypercalcaemia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxia-telangiectasia, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuroprotection |
| 29331825 | NQH 8.3 (Brain- substantia nigra) | | Von Hippel-Lindau (VHL) syndrome , Alzheimer's disease, Stroke, Tuberous sclerosis, hypercalcaemia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxia-telangiectasia, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuroprotection |
| 29331826 | NQH 8.4 (small intestine) | Small intestine | digestive diseases, obesity, diabetes |
| 29331827 | NQH 8.5 (Spinal cord) | Spinal chord | paralysis, neurodegenerative disorders |
| 29331828 | NQH 8.6 (stomach) | Stomach | Stomach cancer |
| 29331830 | NQH 8.7 (Trachea) | Trachea | Airway diseases, infection |
| 87168518 | NQH 9.1 (Sized-MG-63_treatment pool) | | |
| 87168559 | NQH 9.2 (Sized-HEPG2 untreated) | | |
| 35695763 | NQH.10.1 (MCF-7untreated) | Cancer Cell line | Cancer |
| 35695855 | NQH.10.2 (U-937_treatment pool) | Cancer Cell line | Cancer |
| 35695917 | NQH.10.3 (JAR) | Cancer Cell line | Cancer |
| 35696052 | NQH.10.4 (PA-1) | Cancer Cell line | Cancer |
| 35696286 | NQH.10.5 (CADMEC) | Endothelial cells | heart disease, cancer |
| 35696423 | NQH.10.6 (CADMEC_LA) | Endothelial cells | heart disease, cancer |
| 52644045 | NQH.11.1 (SK-PN-DW) | Cancer Cell line | Cancer |
| 52644150 | NQH.11.2 (Chorionic Villus Cells) | Chorionic villus | fertility, birth defects |
| 52644229 | NQH.11.3 (A549) | Cancer Cell line | Cancer |
| 52644296 | NQH.11.4 (U266B1) | Cancer Cell line | Cancer |
| 52644332 | NQH.11.5 (Daoy) | Cancer Cell line | Cancer |
| 52644507 | NQH.11.6 (SW1783) | Cancer Cell line | Cancer |
| 52645080 | NQH.12.1 (U-118MG) | Cancer Cell line | Cancer |
| 52645129 | NQH.12.2 (A204) | Cancer Cell line | Cancer |
| 52645156 | NQH.12.3 (T24) | Cancer Cell line | Cancer |
| 52646317 | NQH.12.4 (G-401) | Cancer Cell line | Cancer |
| 52646365 | NQH.12.5 (CaSki) | Cancer Cell line | Cancer |
| 52646842 | NQH.12.6 (SHP-77) | Cancer Cell line | Cancer |

| | | | |
|----------|---|--------------------------------|--|
| 60424179 | NQH.14.1 (Yale75_breast carcinoma) | Breast carcinoma | Breast Cancer |
| 60424269 | NQH.14.2 (Yale78B_ovarytumor) | Ovary tumor | Ovarian cancer |
| 60431528 | NQH.14.3 (Yale79_prostateBPH) | Prostate | Prostate Cancer |
| 60431602 | NQH.14.4 (Yale80_ProstateAdenocarcinoma) | Prostate | Prostate Cancer |
| 60431735 | NQH.14.5 (Yale86_UterineMyoma) | Uterine Myoma | Uterine Cancer |
| 60431850 | NQH.14.6 (Yale207_Myometrium) | Myometrium | Fertility |
| 60432049 | NQH.15.1 (Yale99_cervix) | Cervix | Osteoporosis, cervical cancer |
| 60432113 | NQH.15.2 (Yale45_spleenITP) | | Hemophilia, Hypercoagulation, Idiopathic thrombocytopenic purpura, Immunodeficiencies, Graft vesus host |
| 60432229 | NQH.15.3 (Yale16_Skin) | Skin | wound healing, melanoma |
| 60432289 | NQH.15.4 (Yale137_Parotid) | | |
| 60433356 | NQH.15.5 (Yale38_SmallIntestine) | Small intestine | digestive diseases, obesity, diabetes |
| 60433438 | NQH.15.6 (Yale28_ColonAscending) | Colon | Colon cancer |
| 65274444 | NQH.17.1 (Larynx) | Larynx | Cancer |
| 65274572 | NQH.17.2 (Duodenum) | Duodenum | |
| 65274620 | NQH.17.3 (Kidney, Primary tumors) | | Diabetes, Autoimmune disease, Renal artery stenosis, Interstitial nephritis, Glomerulonephritis, Polycystic kidney disease, Systemic lupus erythematosus, Renal tubular acidosis, IgA nephropathy, Hypercalceimnia, Lesch-Nyhan syndrome |
| 65274727 | NQH.17.4 (Lung Pleura, normal) | Lung | Airway diseases, infection |
| 65274791 | NQH.17.5 (Lung, Normal Adult) | Lung | Airway diseases, infection |
| 83373044 | NQH.18.230 (Pooled adrenal gland, placenta) | Adrenal Gland/Suprarenal gland | Adrenoleukodystrophy, Congenital Adrenal Hyperplasia, |
| 85658542 | NQH.18.560 (Pooled uterus, BeWo pool) | Uterus | Infertility, birth defects |
| 33656970 | NQH.9.1 (MG-63_treatment pool) | Cancer Cell line | Cancer |
| 33657023 | NQH.9.2 (HEPG2 untreated) | | Von Hippel-Lindau (VHL) syndrome, Cirrhosis, Transplantation |
| 33657084 | NQH.9.3 (PC3_untreated) | Cancer Cell line | Cancer |
| 33657109 | NQH.9.4 (TF-1_TPA) | Cancer Cell line | Cancer |
| 33657182 | NQH.9.5 (TF-1_TPO) | Cancer Cell line | Cancer |
| 33657349 | NQH.9.6 (TF-1_Hemin) | Cancer Cell line | Cancer |
| 33657402 | NQH.9.7 (HFDPC) | Cancer Cell line | Cancer |
| 264259 | NQH1 (Mixture of eight adult & two fetal tissues) | | |
| 264288 | NQH2 (Ten tissues plus lymphocyte control) | | |
| 264448 | NQH3 (Bone Marrow) | Bone Marrow | Hemophilia, hypercoagulation, Idiopathic thrombocytopenic purpura, autoimmune disease, allergies, immunodeficiencies, transplantation, Graft vesus host, |
| 265017 | NQH4.1 (lymph node) | Lymph Node | Lymphedema, Allergies |

| | | | |
|----------|-----------------------------------|--------------------------------|---|
| 265018 | NQH4.2 (fetal kidney) | Fetal Kidney | Diabetes, Autoimmune disease, Renal artery stenosis, Interstitial nephritis, Glomerulonephritis, Polycystic kidney disease, Systemic lupus erythematosus, Renal tubular acidosis, IgA nephropathy, Hypercalcaemia, Lesch-Nyhan syndrome |
| 66712502 | NQH4.2 (Sized) | | |
| 265019 | NQH4.3 (pituitary gland) | | Von Hippel-Lindau (VHL) syndrome, Alzheimer's disease, Stroke, Tuberous sclerosis, hypercalcaemia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxia-telangiectasia, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuroprotection, Obesity |
| 66714117 | NQH4.3 (Sized) | | |
| 265020 | NQH4.4 (testis) | testis | Infertility, birth defects |
| 265021 | NQH4.5 (fetal liver) | Fetal Liver | Von Hippel-Lindau (VHL) syndrome, Cirrhosis, Transplantation |
| 265022 | NQH4.6 (thyroid) | Thyroid | Hyperthyroidism and Hypothyroidism |
| 18108376 | NQH5.1 (MCF-7) | Breast cancer | Breast Cancer |
| 18108387 | NQH5.2 (CCRF-CEM) | Cancer Cell line | Cancer |
| 264952 | NRL1: HPLC FRACTIONATION OF RELIG | | |
| 263971 | Old BB3 Baits | | |
| 263969 | Old BB5 Baits | | |
| 263975 | ORFSEL | | |
| 263972 | OTHER Baits | | |
| 263978 | pGALORF | | |
| 264106 | PPBAITS | | |
| 264088 | QC-YA7 | | |
| 264089 | QC-YA8 | | |
| 264102 | Resequenced Interactors | | |
| 264369 | RRH.1 | | |
| 60170394 | RRH.10.1 (MCF-7untreated) | Breast cancer | Breast Cancer |
| 60170615 | RRH.10.2 (U-937_treatment pool) | Cancer Cell line | Cancer |
| 60170831 | RRH.10.3 (JAR) | Cancer Cell line | Cancer |
| 60174639 | RRH.11.8 (HeLa) | Cancer Cell line | Cancer |
| 264113 | π QEA Baits | | |
| 263973 | RRQEA_B5 baits | | |
| 29146498 | SRD 3.1 (SKMC) | Cancer Cell line | Cancer |
| 29146499 | SRD 3.2 (SKMC) | Cancer Cell line | Cancer |
| 29147620 | SRD 3.3 (RPTEC) | Cancer Cell line | Cancer |
| 29148627 | SRD 3.4 (HRCE) | Cancer Cell line | Cancer |
| 29148629 | SRD 3.6 (HRE) | Cancer Cell line | Cancer |
| 29148784 | SRD 3.7 (HRE) | Cancer Cell line | Cancer |
| 55810764 | SRD.7.1 (Lymph Node) | Lymph Node | Lymphedema, Allergies |
| 55811150 | SRD.7.2 (pancreas) | Pancreas | Pancreatitis, diabetes, pancreatic cancer |
| 55811386 | SRD.7.3 (Adrenal Gland) | Adrenal Gland/Suprarenal gland | Adrenoleukodystrophy, Congenital Adrenal Hyperplasia, |

| | | | |
|----------|---------------------------|--------------------------------|--|
| 55811576 | SRD.7.4 (Pituitary Gland) | Pituitary | Von Hippel-Lindau (VHL) syndrome , Alzheimer's disease, Stroke, Tuberous sclerosis, hypercalcaemia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxia-telangiectasia, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuroprotection, Obesity |
| 55811957 | SRD.7.5 (Fetal Liver) | Fetal Liver | Von Hippel-Lindau (VHL) syndrome, Cirrhosis, Transplantation |
| 55812038 | SRD.7.6 (Fetal Kidney) | Fetal kidney | Diabetes, Autoimmune disease, Renal artery stenosis, Interstitial nephritis, Glomerulonephritis, Polycystic kidney disease, Systemic lupus erythematosus, Renal tubular acidosis, IgA nephropathy, Hypercalcaemia, Lesch-Nyhan syndrome |
| 56181562 | SRD.8.1 (Lymph Node) | Lymph Node | Lymphedema , Allergies |
| 56181686 | SRD.8.2 (Pancreas) | Pancreas | Pancreatitis, diabetes, pancreatic cancer |
| 56182181 | SRD.8.3 (Adrenal Gland) | Adrenal Gland/Suprarenal gland | Adrenoleukodystrophy , Congenital Adrenal Hyperplasia, |
| 56182323 | SRD.8.4 (Pituitary Gland) | Pituitary | Von Hippel-Lindau (VHL) syndrome , Alzheimer's disease, Stroke, Tuberous sclerosis, hypercalcaemia, Parkinson's disease, Huntington's disease, Cerebral palsy, Epilepsy, Lesch-Nyhan syndrome, Multiple sclerosis, Ataxia-telangiectasia, Leukodystrophies, Behavioral disorders, Addiction, Anxiety, Pain, Neuroprotection, Obesity |
| 56182435 | SRD.8.5 (Fetal Liver) | Fetal Liver | Von Hippel-Lindau (VHL) syndrome, Cirrhosis, Transplantation |
| 56182575 | SRD.8.6 (Fetal Kidney) | | Diabetes, Autoimmune disease, Renal artery stenosis, Interstitial nephritis, Glomerulonephritis, Polycystic kidney disease, Systemic lupus erythematosus, Renal tubular acidosis, IgA nephropathy, Hypercalcaemia, Lesch-Nyhan syndrome |
| 32833986 | SRD4: HL adapter | | |
| 56526486 | SRD5: 1:rr fragments | | |
| 33109954 | SRD5: long-RXRJ | | |
| 56994075 | SRD9.1 (CS/SC) | Cancer Cell line | Cancer |
| 263977 | TSC Screen 1 | | |

Disclosed Sequences of ORFX Nucleic Acid and Polypeptide Sequences

<210> 1
 <211> 312
 <212> DNA
 <213> Homo sapiens

<400> 1
 gtcgacagtg tcatggggcgc cgcaccggcg atcggcggtg cgagcacgtt cgggtaccgcg
 60
 agaggaatcg gtaccgcgcg acgaggagga cgggcgggca gctcggcgac gagggtgccg
 120
 gcgagcnncg atccgtctgc ctccccacca ccgatctcgg cagccagggtt gtcgaggatg
 180
 tcgctggtgg agcgtctgct cgcgtctgct tcttgctcag ccgtcttgcg ggactggggc
 240
 tgctcctcgg ccttggtcag ggcagcgtca atgtcggcga aatccgattc tgccttcggc
 300
 tcagagacgc gt
 312

<210> 2
 <211> 104
 <212> PRT
 <213> Homo sapiens

<400> 2
 Val Asp Ser Val Met Gly Ala Arg Pro Ala Ile Gly Val Ala Ser Thr
 1 5 10 15
 Phe Gly Thr Ala Arg Gly Ile Gly Thr Ala Arg Arg Gly Gly Arg Ala
 20 25 30
 Gly Ser Ser Ala Thr Ser Gly Pro Ala Ser Xaa Asp Pro Ser Ala Ser
 35 40 45
 Pro Pro Pro Ile Ser Ala Ala Arg Leu Ser Arg Met Ser Leu Val Glu
 50 55 60
 Arg Cys Ser Ala Ser Ala Ser Cys Ser Ala Val Leu Arg Asp Trp Ala
 65 70 75 80
 Cys Ser Ser Ala Leu Val Arg Ala Ala Ser Met Ser Ala Lys Ser Asp
 85 90 95
 Ser Ala Phe Gly Ser Glu Thr Arg
 100

<210> 3
 <211> 987
 <212> DNA
 <213> Homo sapiens

<400> 3
 gagctcggta cccccgtcgt cgccgtcaag gcgcagggtca aagtgggtgg ccgtggcaag
 60
 gctggcgggtg taaagatcgc taagtccccg gccgaggcgg ctgaacgcgc cgaggccatc
 120
 ctccggtatgg acatcaaggg tcacaccgtt cacaagggtga tggtcgctga aggtgctgac
 180
 attgccgagg aatactactt ctcgatcttg ctcgaccgtg gagagcgtcg ctaccttgcg
 240
 atgtgctcgc gtgagggtgg catggacatc gagacccttg ctaaggagcg ccccgaggct
 300
 ctgcgcaagg tgccggtcga cccgattgac ggcgttgacg atgctaaagc ccgcgaaatc
 360
 cttagcgagg caggggtccc cgactctgag caggacgcta tcgtcccggc tgttctcaag
 420
 ctgtgggaga cctaccgtga cgaggatgcc accctcgtcg aggtcaaccc gatgatcaag
 480
 accggcgatg gacgcatect ggctatcgac ggcaagatga ctgttgacaa caacgcatcc
 540
 ttccgccagc ctgaccgcgc cggcttggtg gatcgcgcca ccaccgaccc gctcgagttg
 600
 cgtgccggcg agctcgggtt caactacgtc aagcttgacg gcaacgtcgg cgtcatcgga
 660
 aacggtgctg gtctggtcat gagcaccctt gactgcgttg cgtacgccgg tgagaacttc
 720
 ccgggatctc cagctccgcg taacttcttc gacatcgggtg gcggcgcttc ggccgagatc
 780
 atggccaacg gccttgacct catcatgagt gacgagcagg tgcgttcctg gttcgtcaac
 840
 gtctttggcg gtatcaccgc ctgtgaccag gtggcgcttg gtatcaaggg cgctctcgag
 900
 aagttgggcg acaaggccgt gaagccgctc gtcgtccgtc tggacggcaa cgctgtggcc
 960
 gagggcagaa agattctcga ggaattc
 987

<210> 4
 <211> 329
 <212> PRT
 <213> Homo sapiens

<400> 4
 Glu Leu Gly Thr Pro Val Val Ala Val Lys Ala Gln Val Lys Val Gly
 1 5 10 15
 Gly Arg Gly Lys Ala Gly Gly Val Lys Ile Ala Lys Ser Pro Ala Glu
 20 25 30
 Ala Ala Glu Arg Ala Glu Ala Ile Leu Gly Met Asp Ile Lys Gly His
 35 40 45

Thr Val His Lys Val Met Val Ala Glu Gly Ala Asp Ile Ala Glu Glu
 50 55 60
 Tyr Tyr Phe Ser Ile Leu Leu Asp Arg Gly Glu Arg Arg Tyr Leu Ala
 65 70 75 80
 Met Cys Ser Arg Glu Gly Gly Met Asp Ile Glu Thr Leu Ala Lys Glu
 85 90 95
 Arg Pro Glu Ala Leu Ala Lys Val Pro Val Asp Pro Ile Asp Gly Val
 100 105 110
 Asp Asp Ala Lys Ala Arg Glu Ile Leu Ser Glu Ala Gly Phe Pro Asp
 115 120 125
 Ser Glu Gln Asp Ala Ile Val Pro Ala Val Leu Lys Leu Trp Glu Thr
 130 135 140
 Tyr Arg Asp Glu Asp Ala Thr Leu Val Glu Val Asn Pro Met Ile Lys
 145 150 155 160
 Thr Gly Asp Gly Arg Ile Leu Ala Ile Asp Gly Lys Met Thr Val Asp
 165 170 175
 Asn Asn Ala Ser Phe Arg Gln Pro Asp Arg Ala Gly Leu Val Asp Arg
 180 185 190
 Ala Thr Thr Asp Pro Leu Glu Leu Arg Ala Gly Glu Leu Gly Leu Asn
 195 200 205
 Tyr Val Lys Leu Asp Gly Asn Val Gly Val Ile Gly Asn Gly Ala Gly
 210 215 220
 Leu Val Met Ser Thr Leu Asp Cys Val Ala Tyr Ala Gly Glu Asn Phe
 225 230 235 240
 Pro Gly Ser Pro Ala Pro Ala Asn Phe Leu Asp Ile Gly Gly Gly Ala
 245 250 255
 Ser Ala Glu Ile Met Ala Asn Gly Leu Asp Leu Ile Met Ser Asp Glu
 260 265 270
 Gln Val Arg Ser Val Phe Val Asn Val Phe Gly Gly Ile Thr Ala Cys
 275 280 285
 Asp Gln Val Ala Leu Gly Ile Lys Gly Ala Leu Glu Lys Leu Gly Asp
 290 295 300
 Lys Ala Val Lys Pro Leu Val Val Arg Leu Asp Gly Asn Ala Val Ala
 305 310 315 320
 Glu Gly Arg Lys Ile Leu Glu Glu Phe
 325

<210> 5

<211> 622

<212> DNA

<213> Homo sapiens

<400> 5

acgcgtggcc tagacctgaa tccattcctc acaaaacagt ctccctcctt gccatgggat
 60
 tccctgggggt ttccacaggc cacagctcta atgggtctgca gcagggttacc ttgttcccca
 120
 gaacatagct tgtcataaca tctctgcagg gttctcccaa acccctttct gcctggcaac
 180
 agctgacatc acacctagct gtaagtcctt gtagatcgca aattactttt tggagactgg
 240
 gggtagcagg ggcattgggg taatagcctt ctagcccttt ttgagggaaa cacatgggtg
 300

aggcattttt ggggctggga agtggggggc tgggtgtccc tggatggctg tgctggcctc
 360
 tggctgcaag ggagaggggc acaggcaagg acatgacccc cgtcaaccct gagccccctc
 420
 cagaaattta accagagcct gtccctcctt tcttgctgc ccccaacatc tcacaatccc
 480
 tctgtgatg gcagatgtct ccatctactc tacagacacc tgcaactatc attcccttga
 540
 tccgtggtaa ttaggagggg actcctctgt gaagaaccgc ttctaccatc ctcttttaga
 600
 aactctttct ccactgggat cc
 622

<210> 6
 <211> 121
 <212> PRT
 <213> Homo sapiens

<400> 6
 Met Ser Leu Pro Val Pro Leu Ser Leu Ala Ala Arg Gly Gln His Ser
 1 5 10 15
 His Pro Gly Asp Thr Arg Pro Pro Leu Pro Ser Pro Lys Ile Ala Ser
 20 25 30
 Pro Met Cys Phe Pro Gln Lys Gly Leu Glu Gly Tyr Tyr Pro Asn Ala
 35 40 45
 Pro Ala Thr Pro Ser Leu Gln Lys Val Ile Cys Asp Leu Gln Gly Leu
 50 55 60
 Thr Ala Arg Cys Asp Val Ser Cys Cys Gln Ala Glu Arg Gly Leu Gly
 65 70 75 80
 Glu Pro Cys Arg Asp Val Met Thr Ser Tyr Val Leu Gly Asn Lys Val
 85 90 95
 Thr Cys Cys Arg Pro Leu Glu Leu Trp Pro Val Lys Thr Pro Gly Asn
 100 105 110
 Pro Met Ala Arg Arg Glu Thr Val Leu
 115 120

<210> 7
 <211> 480
 <212> DNA
 <213> Homo sapiens

<400> 7
 agatcttttc agaagcgggtg ggtggggcgt tcatatccac gtgaatgggc ccaaaaggca
 60
 agaaaaccca cacctctcct cccctcccca gtggctgtgg ctttcctagg gacaatagga
 120
 tgaatgggct ttcagtgtgg ggacagcaaa acatgcacta gggcccagag tggcagttct
 180
 cttggtgtgg agagtgcctg ccacaggcct tggccagagc cctgagggga gtggtgtgtg
 240
 aaaggccacc tccacgtggg taagcgtgag gacttggact tctctggcac tgagatggga
 300
 cctcctgcct gtgggagtca tctggccacc accctggggc cagtaaaggt tggagctaga
 360

agggtcgctcc tccctgactt gagctctgag ggctttgcct gccagccag agcggcaagg
 420
 cacaggggac cctcggggac gcccatggcc accctgggga agacagggt cctcacggt
 480

<210> 8
 <211> 119
 <212> PRT
 <213> Homo sapiens

<400> 8
 Met Gly Phe Gln Cys Gly Asp Ser Lys Thr Cys Thr Arg Ala Gln Ser
 1 5 10 15
 Gly Ser Ser Leu Gly Val Glu Ser Ala Cys His Arg Pro Trp Pro Glu
 20 25 30
 Pro Val Arg Glu Trp Cys Val Lys Gly His Leu His Val Gly Lys Arg
 35 40 45
 Glu Asp Leu Asp Phe Ser Gly Thr Glu Met Gly Pro Pro Ala Cys Gly
 50 55 60
 Ser His Leu Ala Thr Thr Leu Gly Pro Val Lys Val Gly Ala Arg Arg
 65 70 75 80
 Val Val Leu Pro Asp Leu Ser Ser Glu Gly Phe Ala Cys Pro Ala Arg
 85 90 95
 Ala Ala Arg His Arg Gly Pro Ser Gly Thr Pro Met Ala Thr Leu Gly
 100 105 110
 Lys Thr Gly Leu Thr Arg
 115

<210> 9
 <211> 428
 <212> DNA
 <213> Homo sapiens

<400> 9
 tccggaaatt ttccatgaga gactggggca tcgagcagaa gtggatgtct gttctcctgc
 60
 ctctgctgct actttacaat gatccgttct tccccctctc cttcctggtc aacagctggc
 120
 tcccagggat gctggatgac ctctttcagt ccatgttcct gtgcgccctg ctgctcttct
 180
 ggctgtgcgt gtaccacggg attcgtgtcc agggagaaaag aaagtgttta actttctatt
 240
 tgcctaaatt cttcattggt ggactattgt ggttggttc tggtacgcta ggaatatggc
 300
 aaacagttaa cgaattacat gatccaatgt accagtatcg agttgatacc ggaaattttc
 360
 agggaatgaa ggtcttcttc atggtggtgg cagcgggtga cattctgtac ctcttgttct
 420
 tgatagtg
 428

<210> 10
 <211> 138
 <212> PRT

<213> Homo sapiens

<400> 10

```

Met Arg Asp Trp Gly Ile Glu Gln Lys Trp Met Ser Val Leu Leu Pro
 1           5           10           15
Leu Leu Leu Leu Tyr Asn Asp Pro Phe Phe Pro Leu Ser Phe Leu Val
      20           25           30
Asn Ser Trp Leu Pro Gly Met Leu Asp Asp Leu Phe Gln Ser Met Phe
 35           40           45
Leu Cys Ala Leu Leu Leu Phe Trp Leu Cys Val Tyr His Gly Ile Arg
 50           55           60
Val Gln Gly Glu Arg Lys Cys Leu Thr Phe Tyr Leu Pro Lys Phe Phe
 65           70           75           80
Ile Val Gly Leu Leu Trp Leu Ala Ser Val Thr Leu Gly Ile Trp Gln
      85           90           95
Thr Val Asn Glu Leu His Asp Pro Met Tyr Gln Tyr Arg Val Asp Thr
      100           105           110
Gly Asn Phe Gln Gly Met Lys Val Phe Phe Met Val Val Ala Ala Val
      115           120           125
Tyr Ile Leu Tyr Leu Leu Phe Leu Ile Val
      130           135

```

<210> 11

<211> 453

<212> DNA

<213> Homo sapiens

<400> 11

```

cttaagaatc gcctcactca acgggtcagct tgccgaccat gcccgctga taatgccccg
60
gaatgttgca ggcaactca agaccggtgg ccttggtgaa ggtccaggtc agctcgggcg
120
acttgccccg ctgcaccagc acgctggttg ggtcgatcatg cttcatgccg cccatatcgc
180
catgccccat ggcgcggtgg tccatcttgc ccatgccggt ggccgtgagc atgccgctgg
240
cttgcattct gagcatttct ttctggtgtt cggcgtgcat cgccgcatca cccagattga
300
attcgtgcag taactggcct ttgttgacca gcacaaagcg cacggtctca ccggctttta
360
catccagagc cttggggcgaa aaggaaatgt cctgcagggt gacttccacg gtgcgcgtgg
420
ctttatcggc cggtgccggg tggecaaacg cgt
453

```

<210> 12

<211> 130

<212> PRT

<213> Homo sapiens

<400> 12

```

Met Leu Gln Ala Asn Ser Arg Pro Val Ala Leu Val Lys Val Gln Val
 1           5           10           15
Ser Ser Ala Asp Leu Pro Gly Ser Thr Ser Thr Leu Leu Gly Ser Ser

```


| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| | 20 | | 25 | | 30 | | | | | | | | | | |
| Cys | Phe | Met | Pro | Pro | Ile | Ser | Pro | Cys | Pro | Met | Ala | Ala | Trp | Ser | Ile |
| | 35 | | | | | | 40 | | | | | 45 | | | |
| Leu | Pro | Met | Pro | Val | Ala | Val | Ser | Met | Pro | Leu | Ala | Cys | Ile | Leu | Ser |
| | 50 | | | | | 55 | | | | | 60 | | | | |
| Ile | Ser | Phe | Trp | Cys | Ser | Ala | Cys | Ile | Ala | Ala | Ser | Pro | Arg | Leu | Asn |
| 65 | | | | | 70 | | | | 75 | | | | | 80 | |
| Ser | Cys | Ser | Asn | Trp | Pro | Leu | Leu | Thr | Ser | Thr | Lys | Arg | Thr | Val | Ser |
| | | | 85 | | | | | 90 | | | | | | 95 | |
| Pro | Ala | Phe | Thr | Ser | Arg | Ala | Leu | Gly | Glu | Lys | Glu | Met | Ser | Cys | Arg |
| | | | 100 | | | | | 105 | | | | | 110 | | |
| Val | Thr | Ser | Thr | Val | Arg | Val | Ala | Leu | Ser | Ala | Gly | Ala | Gly | Trp | Pro |
| | | 115 | | | | | 120 | | | | | 125 | | | |
| Asn | Ala | | | | | | | | | | | | | | |
| | 130 | | | | | | | | | | | | | | |

<210> 13

<211> 2034

<212> DNA

<213> Homo sapiens

<400> 13

nacgcgttcg gcgtagtccc cttcctgccca tccgtcctga gtcacctgct gcccgctgctg
 60
 ggcggtggcca agcaggacac ggtgcgcgctg gccttctgct ccggggacct gcggctcctc
 120
 tgggccctac catgctggca ttttctccca tgtgtcaaac acatgggttc agccagcgaa
 180
 gattccatgg gacctcctcg tgtgggacgt gtgctcccca ccacaaatgg aacgttcctt
 240
 gtttgcatct ggaggggttg gtggctctgc tggctggagc agcctggggc cagaggaagc
 300
 cgtatcaacc ggctctgcag cgcttcagcg aggggtgccct ggagtaccta gccaacctgg
 360
 accgagcccc agaccccacg gtcaggaagg acgcctttgc caccgacatc ttcagcgctt
 420
 acgatgttct cttccatcag tggctgcaga gtcgagaagc caagctccgt cttgccgtgg
 480
 tggaggctct ggggcctatg agccatctgc tgcccagtga gaggttgga gagcagctgc
 540
 ccaagctcct ccctgggatt ctgcctctct acaagaagca cgcagagacc ttctacttgt
 600
 ccaagagcct gggccagatc ctcgaggcag ctgtgagtgt gggcagccgc aactggaga
 660
 cccagctgga tgccctcttg gctgcactgc actcccagat ctgtgtgcct gtggagtctt
 720
 caagccccct ggtgatgagt aaccagaagg aggtgctgct ctgcttcact gtgctggcct
 780
 gcagctcgcc tgaccgccta ctggccttcc tgctgcccag gctggacacc accaatgaga
 840
 ggacccgcgt gggcaccctg cagggggcca aacatgtcat caactcaact gctgctcaaa
 900
 tggaagataa aaagcccttt atcctgtctt ccatgaggct tcctctcctg aacaccaaca
 960

gcaaggtgaa gcgggcagtg gtgcaggtga ttagcgccca tggcccacca cggctacctg
 1020
 gagcagcctg gaggtgaggc gatgatcgag taaatcgtgc agcagtgcg cgtgcccccc
 1080
 gagcaggagc ctgagaagcc agggccccgc agcaaggacc ccaaggccga cagcgtgcgg
 1140
 gccatcagcg tgcgcaccct ctacctgggc agcaccaccg tggacaggat gagtcacgtc
 1200
 ctctggccat acctgctcca gttcctcacc cctgtgcgct tcaactggggc cctgactccg
 1260
 ctctgcagga gcctcgtgca tctggcgagc aagaggcagg agggccggggc cgacgccttc
 1320
 ctcattccagt acgacgcccc tgcgagcctc ccgtctccct atgctgtaac cggaagactg
 1380
 ttggttgtgt cttccagccc ctacctaggg gacggacgtg gggcagcggc gctgcgctc
 1440
 ctcagtgttc tgcacccaaa cattcaccct ttgctgggtc agcattggga aacgactgtc
 1500
 ccgctgctgc tggggtacct ggatgagcac acagaagaga ccctgccaca ggaggagtgg
 1560
 gaggagaagc tgttgatggt gagggccggg gtacggccca tcctgggcct taagggtgtg
 1620
 tctggcctgg ggggtgctgg ggtggcagag gctggggccac ctgcctcgac ctcacctcgt
 1680
 ggtttggctg gggagccaag gatcaggcag catcaaggct gaagacccca gcagccttgc
 1740
 agcggggggc ttgctgtgac aaggcaccgg ccctctagca gtcgcagccc caagcgctcg
 1800
 gggcaacctc tcacctgcc tggtagacca actgtggcat ggctgtcccc tgagggttgg
 1860
 ctctgccgcc cccggcctcc gctggaaggc ggtctgcagc ccctgcagcc acagcacatg
 1920
 gggatgtgcc caggctccag ccagccctgt gaggggtcgg gctcccagcc cctcagtggc
 1980
 atcttggcct gcagttcctg cgagacaccc tggccatcat ttctgacaac gcgt
 2034

<210> 14

<211> 222

<212> PRT

<213> Homo sapiens

<400> 14

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Ile | Val | Gln | Gln | Cys | Ala | Leu | Pro | Pro | Glu | Gln | Glu | Pro | Glu | Lys | Pro |
| 1 | | | | 5 | | | | | 10 | | | | | 15 | |
| Gly | Pro | Gly | Ser | Lys | Asp | Pro | Lys | Ala | Asp | Ser | Val | Arg | Ala | Ile | Ser |
| | | | 20 | | | | | 25 | | | | 30 | | | |
| Val | Arg | Thr | Leu | Tyr | Leu | Val | Ser | Thr | Thr | Val | Asp | Arg | Met | Ser | His |
| | | 35 | | | | 40 | | | | | 45 | | | | |
| Val | Leu | Trp | Pro | Tyr | Leu | Leu | Gln | Phe | Leu | Thr | Pro | Val | Arg | Phe | Thr |
| | 50 | | | | | 55 | | | | 60 | | | | | |
| Gly | Ala | Leu | Thr | Pro | Leu | Cys | Arg | Ser | Leu | Val | His | Leu | Ala | Gln | Lys |
| 65 | | | | 70 | | | | | 75 | | | | | 80 | |
| Arg | Gln | Glu | Ala | Gly | Ala | Asp | Ala | Phe | Leu | Ile | Gln | Tyr | Asp | Ala | His |

```
<210> 15
<211> 363
<212> DNA
<213> Homo sapiens
```

```
<400> 15
naccgcttgc tggctcgcca cggcaagggc catgtcggct gcgatatctg caagccggcg
60
gtgggttcga tccttgccct gtgctggaac cagccgatca tggacccggc gttggtgccc
120
ttgcaggaca ccaatgacac cttcatggcc aacatgcaga agaacggtac ctattcgatc
180
atcccgcgta tcgccggcgg cgagatcacc ccggacaaac tgatcgccct cggcgcggtg
240
gcgaagaaat acgatctgta caccaagatc accggcggcc agcggatcga cctgttcggc
300
gccagttgc acgaattgcc gcagatctgg ggcgagctgg tggatgccgg attcgagacc
360
ggc
363
```

```
<210> 16
<211> 121
<212> PRT
<213> Homo sapiens
```

```

<400> 16
Xaa Ala Leu Leu Ala Arg His Gly Lys Gly His Val Gly Cys Asp Ile
 1             5             10             15
Cys Lys Pro Ala Val Gly Ser Ile Leu Ala Ser Cys Trp Asn Gln Pro
 20             25             30
Ile Met Asp Pro Ala Leu Val Pro Leu Gln Asp Thr Asn Asp Thr Phe
 35             40             45
Met Ala Asn Met Gln Lys Asn Gly Thr Tyr Ser Ile Ile Pro Arg Ile
 50             55             60
Ala Gly Gly Glu Ile Thr Pro Asp Lys Leu Ile Ala Leu Gly Ala Val

```

```

65              70              75              80
Ala Lys Lys Tyr Asp Leu Tyr Thr Lys Ile Thr Gly Gly Gln Arg Ile
              85              90              95
Asp Leu Phe Gly Ala Gln Leu His Glu Leu Pro Gln Ile Trp Gly Glu
              100              105              110
Leu Val Asp Ala Gly Phe Glu Thr Gly
              115              120

```

```

<210> 17
<211> 682
<212> DNA
<213> Homo sapiens

```

```

<400> 17
gaattccatt ttgtggagta agaggtgact ggggtatagg gtacaacca tagccatcca
60
tgttcatctt tgttttgaat ataattggct agaagatata catatatcta tgtaacttcc
120
tctagcatcc tccagtatgg aggctgcatt aagactgcat gaaggagagg gagagaaggg
180
agaaacagag cagctggaca agaggacagg tataggggaat aagggagaag ccagtaaggc
240
aggaaagacc ctccgtgaca aaggggcagg gaacagaact caaacattta atggcaggta
300
accaggtta gaatggtaaa ttgaaaggtg aatataaagg gagaatggtg aaatgaattt
360
tctgaaatta attgctgtgt ttatagtttt tagccatgca tcggaatcac ctcaggactc
420
cactcccaat caattatata tctgggggag gaccaaggcg ttggtatttt tcagaagctc
480
cactggtgat tctgacagca cagctaggat taagaaactg atcaatggga acagcatgcc
540
tgttgcagag gagcttcctt gggaaatgtc acacacagaa catcaatctt ccttccccac
600
tcctgagatc cctcattctt tggcaccagg aacagttgca attagtaaac cctggttccc
660
tgctgtctca caaatcgcaa ga
682

```

```

<210> 18
<211> 110
<212> PRT
<213> Homo sapiens

```

```

<400> 18
Met Asn Phe Leu Lys Leu Ile Ala Val Phe Ile Val Phe Ser His Ala
1              5              10              15
Ser Glu Ser Pro Gln Asp Ser Thr Pro Asn Gln Leu Tyr Ile Trp Gly
              20              25              30
Arg Thr Lys Ala Leu Val Phe Phe Arg Ser Ser Thr Gly Asp Ser Asp
              35              40              45
Ser Thr Ala Arg Ile Lys Lys Leu Ile Asn Gly Asn Ser Met Pro Val
50              55              60
Ala Glu Glu Leu Pro Trp Glu Met Ser His Thr Glu His Gln Ser Ser

```

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 65 | | 70 | | 75 | | 80 | | | | | | | | | |
| Phe | Pro | Thr | Pro | Glu | Ile | Pro | His | Ser | Leu | Ala | Pro | Gly | Thr | Val | Ala |
| | | | 85 | | | | | | 90 | | | | | 95 | |
| Ile | Ser | Lys | Pro | Trp | Phe | Pro | Ala | Val | Ser | Gln | Ile | Ala | Arg | | |
| | | | 100 | | | | | | 105 | | | | | 110 | |

<210> 19

<211> 515

<212> DNA

<213> Homo sapiens

<400> 19

```

cttggctggc agacatggga cctgcttccc tcttacaccc cagtcttggc aaggatcatg
60
cccccatctc aactatgtta gccagtctgg ctgttcactt agtcactaca gtttgcttct
120
cgtctgcagt gcagtcttgg gctataagaa aactggggcc actcaatacc tccccctttt
180
tgcccttctt cctcctctgg tccatgggtg ggggtggggg gagcccagtt tcagcaccag
240
cagctggagc ccataccaca ctcatTTTTc agttctggct gtgggagccc ctcccacagg
300
tttcagttcc ccaagcccca ggctgagtt ttttttattg caaaagctgg ttgttgttgt
360
ggctagctcc caggcgtgtg aggtgcagct tgctaagtaa gagctaggaa agagaatagg
420
gtcctgctgt aggtgtccag tctgaaggaa tgcttgggat acttcctcaa gcagttcctt
480
ctcacagtct cctggctgct ccgcatgtca gatct
515

```

<210> 20

<211> 130

<212> PRT

<213> Homo sapiens

<400> 20

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Gly | Pro | Ala | Ser | Leu | Leu | His | Pro | Ser | Leu | Gly | Lys | Asp | His | Ala |
| 1 | | | | 5 | | | | 10 | | | | | | 15 | |
| Pro | Ile | Ser | Thr | Met | Leu | Ala | Ser | Leu | Ala | Val | His | Leu | Val | Thr | Thr |
| | | | 20 | | | | | 25 | | | | | | 30 | |
| Val | Cys | Phe | Ser | Ser | Ala | Val | Gln | Ser | Trp | Ala | Ile | Arg | Asn | Thr | Gly |
| | 35 | | | | | | 40 | | | | | 45 | | | |
| Pro | Leu | Asn | Thr | Ser | Pro | Leu | Leu | Ala | Leu | Leu | Leu | Leu | Trp | Ser | Met |
| | 50 | | | | | 55 | | | | | 60 | | | | |
| Gly | Gly | Val | Gly | Gly | Ser | Pro | Val | Ser | Ala | Pro | Ala | Ala | Gly | Ala | His |
| 65 | | | | | 70 | | | | | 75 | | | | 80 | |
| Thr | Thr | Leu | Ile | Phe | Gln | Phe | Trp | Leu | Trp | Glu | Pro | Leu | Pro | Gln | Val |
| | | | 85 | | | | | 90 | | | | | | 95 | |
| Ser | Val | Pro | Gln | Ala | Pro | Gly | Leu | Ser | Phe | Phe | Tyr | Cys | Lys | Ser | Trp |
| | | | 100 | | | | | 105 | | | | | | 110 | |
| Leu | Leu | Leu | Trp | Leu | Ala | Pro | Arg | Val | Arg | Cys | Ser | Leu | Leu | Ser | |
| | | | 115 | | | | 120 | | | | | | | 125 | |
| Lys | Ser | | | | | | | | | | | | | | |

130

<210> 21
 <211> 390
 <212> DNA
 <213> Homo sapiens

<400> 21
 gtgcgacaaa aagagcacgt tcgcaagggg aggaagagcg tgccaccggt tctgccgagc
 60
 tagacgcggt gcctatgggt gcggaggacc atggagtga gcgagtaaga ctagatgatg
 120
 caacaaatgt gcctgagggg gaaatggcac gagccagtgc caatgagggc atgacacctg
 180
 ttaaccacga caaataccct tctgtccttt taaatgaagc ggcccaggct tcattactgg
 240
 atacaatgac tgcttgact gatgggttca caattgagca attggagctt acacgatctc
 300
 tatgttatga aagagtatta gcacatcgat cctcatggga tcgttcagcc ctgggtcaag
 360
 aattaaagca agttgtccaa ggcatccatn
 390

<210> 22
 <211> 105
 <212> PRT
 <213> Homo sapiens

<400> 22
 Met Val Ala Glu Asp His Gly Val Lys Arg Val Arg Leu Asp Asp Ala
 1 5 10 15
 Thr Asn Val Pro Glu Gly Glu Met Ala Arg Ala Ser Ala Asn Glu Gly
 20 25 30
 Met Thr Pro Val Asn His Asp Lys Tyr Pro Ser Val Leu Leu Asn Glu
 35 40 45
 Ala Ala Gln Ala Ser Leu Leu Asp Thr Met Thr Ala Cys Thr Asp Gly
 50 55 60
 Phe Thr Ile Glu Gln Leu Glu Leu Thr Arg Ser Leu Cys Tyr Glu Arg
 65 70 75 80
 Val Leu Ala His Arg Ser Ser Trp Asp Arg Ser Ala Leu Ala Gln Glu
 85 90 95
 Leu Lys Gln Val Val Gln Gly Ile His
 100 105

<210> 23
 <211> 385
 <212> DNA
 <213> Homo sapiens

<400> 23
 ntctcggagg ccgacagcct ggcgggctgg aagccctcgg tgtaccacgt gctgctcatc
 60
 ctgggcctgt tcgccgtgct gctgtcctgc tgcgcctcgg ccatgtacac cagcgtggag
 120

ggctgggact acgtggactc gctctacttc tgettctgtca ccttcagcac catcggtctc
 180
 ggggacctgg tgagcagcca gcacgcgcc taccggaacc aggggtctta ccgcctgggc
 240
 aacttcctct tcatcctgct cggcgtgtgc tgcatttact cgctcttcaa cgtcatctcc
 300
 atcctcatca agcagggtgct caactggatg ctgcgcaagc tgagctgccg ctgctgcgcg
 360
 cgctgctgcc cggctcctgg cgcgc
 385

<210> 24

<211> 128

<212> PRT

<213> Homo sapiens

<400> 24

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Xaa | Ser | Glu | Ala | Asp | Ser | Leu | Ala | Gly | Trp | Lys | Pro | Ser | Val | Tyr | His |
| 1 | | | | 5 | | | | 10 | | | | | 15 | | |
| Val | Leu | Leu | Ile | Leu | Gly | Leu | Phe | Ala | Val | Leu | Leu | Ser | Cys | Cys | Ala |
| | | | 20 | | | | | 25 | | | | | 30 | | |
| Ser | Ala | Met | Tyr | Thr | Ser | Val | Glu | Gly | Trp | Asp | Tyr | Val | Asp | Ser | Leu |
| | | 35 | | | | | 40 | | | | | 45 | | | |
| Tyr | Phe | Cys | Phe | Val | Thr | Phe | Ser | Thr | Ile | Gly | Phe | Gly | Asp | Leu | Val |
| | | 50 | | | | 55 | | | | | 60 | | | | |
| Ser | Ser | Gln | His | Ala | Ala | Tyr | Arg | Asn | Gln | Gly | Leu | Tyr | Arg | Leu | Gly |
| 65 | | | | 70 | | | | | 75 | | | | | 80 | |
| Asn | Phe | Leu | Phe | Ile | Leu | Leu | Gly | Val | Cys | Cys | Ile | Tyr | Ser | Leu | Phe |
| | | | | 85 | | | | 90 | | | | | | 95 | |
| Asn | Val | Ile | Ser | Ile | Leu | Ile | Lys | Gln | Val | Leu | Asn | Trp | Met | Leu | Arg |
| | | | 100 | | | | | 105 | | | | | 110 | | |
| Lys | Leu | Ser | Cys | Arg | Cys | Cys | Ala | Arg | Cys | Cys | Pro | Ala | Pro | Gly | Ala |
| | | 115 | | | | | 120 | | | | | 125 | | | |

<210> 25

<211> 337

<212> DNA

<213> Homo sapiens

<400> 25

ccatgggaga gaccgtgcat tttcttctag gtctgcgtgg gaagtcactg cagagtttctg
 60
 aggaggggag ttcccagctc tgtatttttg aagggtcagt cttgttgctt ggaccagtga
 120
 ggagccccgt gggatccaga ctcgagtggg tggagccggg gcaggtggga gcagagacac
 180
 tggaggaaaag ctggtcgaat gcactgtgta tttggaggca gaaccagcag agggtcctct
 240
 gggttgagtg tagggcaaaa gagaaagaag gcaccaagcc tggggctctgg gttttctctc
 300
 ttacacttgc tgggtggacg gtggtgccac tgaatga
 337

<210> 26

<211> 111
 <212> PRT
 <213> Homo sapiens

<400> 26
 Met Gly Glu Thr Val His Phe Leu Leu Gly Leu Arg Gly Lys Ser Leu
 1 5 10 15
 Gln Ser Phe Glu Glu Gly Ser Ser Gln Leu Cys Ile Phe Glu Gly Ser
 20 25 30
 Val Leu Leu Leu Gly Pro Val Arg Ser Pro Val Gly Ser Arg Leu Glu
 35 40 45
 Trp Val Glu Pro Gly Gln Val Gly Ala Glu Thr Leu Glu Glu Ser Trp
 50 55 60
 Ser Asn Ala Leu Cys Ile Trp Arg Gln Asn Gln Gln Arg Val Leu Trp
 65 70 75 80
 Val Glu Cys Arg Ala Lys Glu Lys Glu Gly Thr Lys Pro Gly Val Trp
 85 90 95
 Val Phe Ser Leu Thr Leu Ala Gly Trp Thr Val Val Pro Leu Asn
 100 105 110

<210> 27
 <211> 333
 <212> DNA
 <213> Homo sapiens

<400> 27
 ccgacgtcga atatccatgc agccgcgcgcg aggatggaga gagcgatgga gcaactcaac
 60
 cgccctgacgc gctcgtgctgcg ccgcgcgcgc accgtggagt tgcccagga taatgaaact
 120
 gctgtttata cattaatgcc aatggttatg gctgatcaac acaggtctgt ttctgaacta
 180
 ctatcaaatt caaaatttga tgtcaattat gcattcggac gtgtgaaaag aagcttgctt
 240
 cacattgcag caaattgtgg atcgggtggaa tgcttggttt tgctgttaaa gaaaggagca
 300
 aatcctaact atcaagatat ttcaggctgt aca
 333

<210> 28
 <211> 111
 <212> PRT
 <213> Homo sapiens

<400> 28
 Pro Thr Ser Asn Ile His Ala Ala Ala Pro Arg Met Glu Arg Ala Met
 1 5 10 15
 Glu Gln Leu Asn Arg Leu Thr Arg Ser Leu Arg Arg Ala Arg Thr Val
 20 25 30
 Glu Leu Pro Glu Asp Asn Glu Thr Ala Val Tyr Thr Leu Met Pro Met
 35 40 45
 Val Met Ala Asp Gln His Arg Ser Val Ser Glu Leu Ser Asn Ser
 50 55 60
 Lys Phe Asp Val Asn Tyr Ala Phe Gly Arg Val Lys Arg Ser Leu Leu

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 65 | | 70 | | 75 | | 80 | | | | | | | | | |
| His | Ile | Ala | Ala | Asn | Cys | Gly | Ser | Val | Glu | Cys | Leu | Val | Leu | Leu | Leu |
| | | 85 | | 90 | | 95 | | | | | | | | | |
| Lys | Lys | Gly | Ala | Asn | Pro | Asn | Tyr | Gln | Asp | Ile | Ser | Gly | Cys | Thr | |
| | | 100 | | 105 | | 110 | | | | | | | | | |

<210> 29
 <211> 375
 <212> DNA
 <213> Homo sapiens

<400> 29
 ncgccgtccg tgctggetat tatgacggcg ggtagcgacc agggcgagga ggtcaactcg
 60
 gagagctatt tgagcgccgt gacgccgtg agtcccaaag agattcgtca gctgccccgc
 120
 tacaatatca cgatcaagcg cgtcgtgaac atgacgggca agggccgcac gccgagctgg
 180
 tactcgctcg tcgtggctgg caatggctcg ggcctcgtgg gctatggcga aggcaaagat
 240
 actaacatca gccgcgcgaa caaaaaggcg ttccacgccg cggtgaaaaa catggacttg
 300
 gtatcggtcc accggtcgaa gagtggcgcc aacacgctcg agccccccgt cgagggccgc
 360
 tggggcgcta cgcgt
 375

<210> 30
 <211> 125
 <212> PRT
 <213> Homo sapiens

| |
|---|
| <400> 30 |
| Xaa Pro Ser Val Leu Ala Ile Met Thr Ala Gly Ser Asp Gln Gly Glu |
| 1 5 10 15 |
| Glu Val Asn Ser Glu Ser Tyr Leu Ser Ala Val Thr Pro Leu Ser Pro |
| 20 25 30 |
| Lys Glu Ile Arg Gln Leu Pro Arg Tyr Asn Ile Thr Ile Lys Arg Val |
| 35 40 45 |
| Val Asn Met Thr Gly Lys Gly Arg Thr Pro Ser Trp Tyr Ser Leu Val |
| 50 55 60 |
| Val Ala Gly Asn Gly Arg Gly Leu Val Gly Tyr Gly Glu Gly Lys Asp |
| 65 70 75 80 |
| Thr Asn Ile Ser Arg Ala Asn Lys Lys Ala Phe His Ala Ala Val Lys |
| 85 90 95 |
| Asn Met Asp Leu Val Ser Val His Arg Ser Lys Ser Gly Ala Asn Thr |
| 100 105 110 |
| Leu Glu Pro Pro Val Glu Gly Arg Trp Gly Ala Thr Arg |
| 115 120 125 |

<210> 31
 <211> 375
 <212> DNA
 <213> Homo sapiens

<400> 31
 accggtcttg gcctcagctt tgctctgaaa ttgaagtcgg tgccaaaagt ggggaagagc
 60
 gggagcaggc acttacgagc ctgcgcgtca gggatgcttc ctggggcccct gagagtgcag
 120
 agattcctgg atccagagct gcggctgggc ggctgcagct gcgcctggga gtgcagggtc
 180
 cccgccctgc cagctcaaaa ggaaatgggg gtcctgcct gttcctggct cctgttggcc
 240
 ctgcagagtg cacaaacctt gccgcgcttc ctccactgca gcttacgtct ttgcagcagc
 300
 cactcccgat gggctgccac tgccatctgt gagaccataa tgtgtgcaat ttgagactca
 360
 tggcctgcat tgttt
 375

<210> 32
 <211> 118
 <212> PRT
 <213> Homo sapiens

<400> 32
 Met Gln Ala Met Ser Leu Lys Leu His Thr Leu Trp Ser His Arg Trp
 1 5 10 15
 Gln Trp Gln Pro Ile Gly Ser Gly Cys Lys Asp Val Ser Cys Ser
 20 25 30
 Gly Gly Ser Ala Ala Arg Phe Val His Ser Ala Gly Pro Thr Gly Ala
 35 40 45
 Arg Asn Arg Gln Glu Pro Pro Phe Pro Phe Glu Leu Ala Gly Arg Glu
 50 55 60
 Pro Cys Thr Pro Arg Arg Ser Cys Ser Arg Pro Ala Ala Ala Leu Asp
 65 70 75 80
 Pro Gly Ile Ser Ala Leu Ser Gly Ala Gln Glu Ala Ser Leu Thr Arg
 85 90 95
 Arg Leu Val Ser Ala Cys Ser Arg Ser Ser Pro Leu Leu Ala Pro Thr
 100 105 110
 Ser Ile Ser Glu Gln Ser
 115

<210> 33
 <211> 351
 <212> DNA
 <213> Homo sapiens

<400> 33
 ccatgcagcc caaccgttgg cgataaagtc cgtttaggcg ataccaattt atgggcaacc
 60
 attgaacaag atttattaac caaagggtgat gagtgtaaat ttgggtggcgg taaaagtgtg
 120
 cgtgatggta tggcgcaaag cggcaccgca actcgcgaca atccaaatgt attggatttt
 180
 gtgattacca atgtgatgat cattgatgcc aaattaggca ttatcaaage cgatattggt
 240

attcgcgatg gtcgtattgt cggatcggga caagcaggta accctgacac catggatgac
 300
 gtcacgccaa acatgattat cggtgctagc acagaagtac ataacgggtgc a
 351

<210> 34
 <211> 117
 <212> PRT
 <213> Homo sapiens

<400> 34
 Pro Cys Ser Pro Thr Val Gly Asp Lys Val Arg Leu Gly Asp Thr Asn
 1 5 10 15
 Leu Trp Ala Thr Ile Glu Gln Asp Leu Leu Thr Lys Gly Asp Glu Cys
 20 25 30
 Lys Phe Gly Gly Gly Lys Ser Val Arg Asp Gly Met Ala Gln Ser Gly
 35 40 45
 Thr Ala Thr Arg Asp Asn Pro Asn Val Leu Asp Phe Val Ile Thr Asn
 50 55 60
 Val Met Ile Ile Asp Ala Lys Leu Gly Ile Ile Lys Ala Asp Ile Gly
 65 70 75 80
 Ile Arg Asp Gly Arg Ile Val Gly Ile Gly Gln Ala Gly Asn Pro Asp
 85 90 95
 Thr Met Asp Asp Val Thr Pro Asn Met Ile Ile Gly Ala Ser Thr Glu
 100 105 110
 Val His Asn Gly Ala
 115

<210> 35
 <211> 355
 <212> DNA
 <213> Homo sapiens

<400> 35
 nngctagctg caccaccacc tgttcattgca ggcagagcgg ccaccctca tggaagaaga
 60
 ggaatccact gtattgggca caggcttctt gctggacctt ggcaagcagg tgcttggtg
 120
 gtaccaggaa gtccagcgtg tacctcagtg cgtcctcccg ataagtcctc tccaccacct
 180
 ggaacacctg gcccaacagg gtgggggctg ttgcctcaaa gggaggatac agggcggcga
 240
 gagggtcttg cacacagtcc tccactggct caggctccat ggctcggcgc cgggcccgcgt
 300
 ccgacgcttg gtcgggcggg cggggccggg cgcgccaccg cctcccttca cgcgt
 355

<210> 36
 <211> 118
 <212> PRT
 <213> Homo sapiens

<400> 36
 Xaa Leu Ala Ala Pro Pro Pro Val His Ala Gly Arg Ala Ala Thr Pro

```

      1           5           10           15
His Gly Arg Arg Gly Ile His Cys Ile Gly His Arg Leu Pro Ala Gly
      20           25           30
Pro Trp Gln Ala Gly Ala Trp Leu Val Pro Gly Ser Pro Ala Cys Thr
      35           40           45
Ser Val Arg Pro Pro Asp Lys Ser Ser Pro Pro Pro Gly Thr Pro Gly
      50           55           60
Pro Thr Gly Trp Gly Leu Leu Pro Gln Arg Val Asp Thr Gly Arg Arg
      65           70           75           80
Glu Cys Ser Ala His Ser Pro Pro Leu Ala Gln Ala Pro Trp Leu Gly
      85           90           95
Ala Gly Pro Arg Pro Thr Leu Gly Arg Ala Gly Gly Ala Gly Arg Ala
      100           105           110
Thr Ala Ser Leu His Ala
      115

```

<210> 37
 <211> 492
 <212> DNA
 <213> Homo sapiens

```

<400> 37
acgcgtggcc ttcgtctgcc accaggaccg actcagcccc accgggtttc cggaccgcgc
60
gcaaccatga caagggcgat gttgtgatct ggggtggattc cttctccgac atgctcgagg
120
gatecggatct ctccggcgga gtcacgggtgc ttgccgaggc cggctatcgc ccacgggtcc
180
tcgccgacga cgtctgctgc ggggtgacgt ggatcactac cggtcagctc gacggtgctc
240
ggcgctcggt ggcgctggt ctcgacgtgc tggcaccctc gtcagacgcc agcgtcccag
300
tcgttgggct agagccgtcc tgcactaccg tctggcggtga tgacgcactc cgcctcctgc
360
cagatgatcc gcgcgtccac cgggtagcca gaaacatgca taccgtcgcc gagatgcttg
420
aggcagcaca gtggacccca ccctcgctag caggccacac cctcgctcgt cagccccatt
480
gtcatcccgc gg
492

```

<210> 38
 <211> 127
 <212> PRT
 <213> Homo sapiens

```

<400> 38
Met Leu Glu Gly Ser Asp Leu Ser Ala Val Val Thr Val Leu Ala Glu
      1           5           10           15
Ala Gly Tyr Arg Pro Arg Val Leu Ala Asp Asp Val Cys Cys Gly Leu
      20           25           30
Thr Trp Ile Thr Thr Gly Gln Leu Asp Gly Ala Arg Arg Arg Leu Arg
      35           40           45
Ala Gly Leu Asp Val Leu Ala Pro Leu Ser Asp Ala Ser Val Pro Val

```

```

      50              55              60
Val Gly Leu Glu Pro Ser Cys Thr Thr Val Trp Arg Asp Asp Ala Leu
65              70              75              80
Arg Leu Leu Pro Asp Asp Pro Arg Val His Arg Val Ala Arg Asn Met
      85              90              95
His Thr Val Ala Glu Met Leu Glu Ala Ala Gln Trp Thr Pro Pro Ser
      100             105             110
Leu Ala Gly His Thr Leu Val Ala Gln Pro His Cys His Pro Ala
      115             120             125

```

<210> 39
 <211> 412
 <212> DNA
 <213> Homo sapiens

```

<400> 39
aacgaaggtn ccgtacgcgc tctgaaagcc ctgcgtaaaag agcggtccga tcgccgggaa
60
gtgatngca ccgcaaaat gcaggtggtc gaagccgcga gttcaggcaa gattgtcttt
120
gaaatggaag acgtttatta cagcattgcc ggaaaacaac tggtagagcaa cttctctgcg
180
caagtcatgc gtggtgataa aattgcgctg attggcccga acggttgtgg taaaacgacg
240
ttgctgaaac tgatgttaag taagattcag gcagacagcg gccgtgttca ctgcggtact
300
aaactggaag ttgcgtactt cgaccagcac cgtgctgagc tggatcctga gcgtacggtg
360
atggataacc tggccgaagg taagcaggaa gtgatggtaa atggccgtgt an
412

```

<210> 40
 <211> 137
 <212> PRT
 <213> Homo sapiens

```

<400> 40
Asn Glu Gly Xaa Val Arg Ala Leu Lys Ala Leu Arg Lys Glu Arg Ser
1              5              10              15
Asp Arg Arg Glu Val Met Xaa Thr Ala Lys Met Gln Val Val Glu Ala
20             25             30
Ala Ser Ser Gly Lys Ile Val Phe Glu Met Glu Asp Val Tyr Tyr Ser
35             40             45
Ile Ala Gly Lys Gln Leu Val Ser Asn Phe Ser Ala Gln Val Met Arg
50             55             60
Gly Asp Lys Ile Ala Leu Ile Gly Pro Asn Gly Cys Gly Lys Thr Thr
65             70             75             80
Leu Leu Lys Leu Met Leu Ser Lys Ile Gln Ala Asp Ser Gly Arg Val
85             90             95
His Cys Gly Thr Lys Leu Glu Val Ala Tyr Phe Asp Gln His Arg Ala
100            105            110
Glu Leu Asp Pro Glu Arg Thr Val Met Asp Asn Leu Ala Glu Gly Lys
115            120            125
Gln Glu Val Met Val Asn Gly Arg Val

```

130

135

<210> 41
 <211> 1080
 <212> DNA
 <213> Homo sapiens

<400> 41
 gaattcaagt ggacacaggc tccacgcccg cgtctcaccg ataagagcta caagcacaac
 60
 tactatgacg agcgggtttc gtcgaagag cgtcttgagc gcactgtggc taaggatttc
 120
 gtcacgacgg aggtcgagcc catgtgggat gcggctgatg tcatgcggat gggtaaggat
 180
 ctcttcatcc agcacggtct gacgacaaat cggaagtcaa tggagtgggt taagcgttac
 240
 taccgccgatt tccgcgttca cgcggtgaat ttccctgggg atccgtaccc gatccatc
 300
 gacgcgacct ttgtgccgct tcgtccgggg ctcatcatca acaacccgaa tcgtccactg
 360
 ccgcaggagc agaggaagat cttcgaggcc aatgactggc agatcggtga tgctgctcag
 420
 ccggcgcacg acacgcctcc agaattgtgc tactcgtctg tgtggctatc aatgaactgc
 480
 ttggtacttg atccgaagac ggtcatctgc gaggcttcgg aagttcatca gatggagcag
 540
 atggacaagc tgggtatgaa cgtcatcccg gtcgccttcc gtgacgcgta cccattcggg
 600
 ggaggtctcc actgcgccac agctgatgta tatcgccaag gtacctgtga ggactacttc
 660
 ccgaatcagg tcgacgaccc gaccttggtg tgagaaaacc ccgtgggtcat gtcatgactg
 720
 acggatctcg gtggctcggg acggaactta cgttgctcgt taccgggccg ccgggtctga
 780
 tatggcagta tcacgcctag caaaaaggag catgtcatgg acatggagcc gggcatcatc
 840
 aacgtcaaac aggaagtctc aggcgtcggg acgatgaacc agaaagtggg attcgtgtcc
 900
 atgcttcttt ctgcaacggg tatggggttg gtgggtactt tcgggcgtct cagcactcct
 960
 gtggatccca cgacgggcag taagtacatc atcgggtgatt ttttggccac tggtaggatg
 1020
 atagtcgggg tcctgggatt tctgcttatt atcgtcatatc ttggaaaatg gtctgagctc
 1080

<210> 42
 <211> 230
 <212> PRT
 <213> Homo sapiens

<400> 42
 Glu Phe Lys Trp Thr Gln Ala Pro Arg Pro Arg Leu Thr Asp Lys Ser
 1 5 10 15
 Tyr Lys His Asn Tyr Tyr Asp Glu Arg Val Ser Leu Glu Glu Arg Leu

20 25 30
 Glu Arg Thr Val Ala Lys Asp Phe Val Thr Thr Glu Val Glu Pro Met
 35 40 45
 Trp Asp Ala Ala Asp Val Met Arg Met Gly Lys Asp Leu Phe Ile Gln
 50 55 60
 His Gly Leu Thr Thr Asn Arg Lys Ser Met Glu Trp Phe Lys Arg Tyr
 65 70 75 80
 Tyr Pro Asp Phe Arg Val His Ala Val Asn Phe Pro Gly Asp Pro Tyr
 85 90 95
 Pro Ile His Ile Asp Ala Thr Phe Val Pro Leu Arg Pro Gly Leu Ile
 100 105 110
 Ile Asn Asn Pro Asn Arg Pro Leu Pro Gln Glu Gln Arg Lys Ile Phe
 115 120 125
 Glu Ala Asn Asp Trp Gln Ile Val Asp Ala Ala Gln Pro Ala His Asp
 130 135 140
 Thr Pro Pro Glu Leu Cys Tyr Ser Ser Val Trp Leu Ser Met Asn Cys
 145 150 155 160
 Leu Val Leu Asp Pro Lys Thr Val Ile Cys Glu Ala Ser Glu Val His
 165 170 175
 Gln Met Glu Gln Met Asp Lys Leu Gly Met Asn Val Ile Pro Val Ala
 180 185 190
 Phe Arg Asp Ala Tyr Pro Phe Gly Gly Gly Leu His Cys Ala Thr Ala
 195 200 205
 Asp Val Tyr Arg Glu Gly Thr Cys Glu Asp Tyr Phe Pro Asn Gln Val
 210 215 220
 Asp Asp Pro Thr Leu Val
 225 230

<210> 43
 <211> 358
 <212> DNA
 <213> Homo sapiens

<400> 43
 gggcccccca catagtggac acaggtttct gggatgtcag catggagtgc caagaggtgg
 60
 gtgaccacct ggtggggaat aaggcgcttc tgggacatag aggctgcctt ccagctgcgc
 120
 ctggcagagc tgttgacaca acagcatggt ctgcagtgcc gggccactgc cacgcacacc
 180
 gatgtccttt aaggatggat ttgggttttc ggattcgcgt ggcctatcag cgggagtccc
 240
 agatcctgaa ggaagtgcag agcccagagg ggatgatctc gctgaggagc acagctgcct
 300
 ccctccgcct tgagagagac acaaggcagt tgccactgct caccagtgcc ctgcacgn
 358

<210> 44
 <211> 105
 <212> PRT
 <213> Homo sapiens

<400> 44
 Met Glu Cys Gln Glu Val Gly Asp His Leu Val Gly Asn Lys Ala Leu

| | | | |
|---|-----|----|----|
| 1 | 5 | 10 | 15 |
| Leu Gly His Arg Gly Cys Leu Pro Ala Ala Pro Gly Arg Ala Val Asp | | | |
| 20 | 25 | 30 | |
| Thr Thr Ala Trp Ser Ala Val Pro Gly His Cys His Ala His Arg Cys | | | |
| 35 | 40 | 45 | |
| Pro Leu Arg Met Asp Leu Gly Phe Arg Ile Arg Val Ala Tyr Gln Arg | | | |
| 50 | 55 | 60 | |
| Glu Ser Gln Ile Leu Lys Glu Val Gln Ser Pro Glu Gly Met Ile Ser | | | |
| 65 | 70 | 75 | 80 |
| Leu Arg Asp Thr Ala Ala Ser Leu Arg Leu Glu Arg Asp Thr Arg Gln | | | |
| 85 | 90 | 95 | |
| Leu Pro Leu Leu Thr Ser Ala Leu His | | | |
| 100 | 105 | | |

<210> 45

<211> 905

<212> DNA

<213> Homo sapiens

<400> 45

gtcgacgata aaggagtatt tgcgcagcag cagtatgatg ctctcgttga ggccgggtttc
 60
 gcggctcctg gaatcccaga gcagtatggg ggcgacgggtg cggatgcatg tgcgtccgca
 120
 ataatcatgg aagaggtcgc tcgagtcctgt gcgtcgtcgt ccaccgtcat atcgtccaat
 180
 gagcttggtg ccgctccctct cctcaaatac ggtagcgagg agcagaggaa acgttatctt
 240
 tctgaagttg cttcgggtaa ggcacttttc ggatatgcgc tctccgaggc tgatgctgga
 300
 tcagatccag ctgcacttaa gtgtcgagcc gacgaagatg gggacagttt cgtcctgaat
 360
 ggcgttaagg cttgggtcac ggaggctggc gaggccaagt acctggtgat atttgcggtt
 420
 actgaccag acgatccgcg ccacagaatc agcgcgttga tggatccatgc agatgaccg
 480
 ggcattagct acggggctcc ggagcacaaa atggggatac gcgggtcagt taccagggaa
 540
 gtggttttca agaatacgcg tatccccaag gaacgagtaa ttggccgctg agggcacggt
 600
 ctgagtgttg ctctaggtac gcttgataac tctcgtgtct cgattgctgc tcaagcagt
 660
 ggaattgccc aaggagcttt agacattgcc acggattacg tccagaagcg caagcagttt
 720
 ggccagccac tgtccaattt tgagggaatc cagttcatgc tcgcagacat ggcaatgcgt
 780
 ttggaggcgg cgcgagcgt gacatactct gcagctgacg gtagtgggcg ccagactgac
 840
 gatgtgagtt acttcggcgc ggcggccaaa tgtttcgctt ccgacacagc gatggcagtg
 900
 tgcac
 905

<210> 46

<211> 301
 <212> PRT
 <213> Homo sapiens

<400> 46
 Val Asp Asp Lys Gly Val Phe Ala Gln Gln Gln Tyr Asp Ala Leu Val
 1 5 10 15
 Glu Ala Gly Phe Ala Ala Pro Gly Ile Pro Glu Gln Tyr Gly Gly Asp
 20 25 30
 Gly Ala Asp Ala Ile Ala Ser Ala Ile Ile Met Glu Glu Val Ala Arg
 35 40 45
 Val Cys Ala Ser Ser Ser Thr Val Ile Ser Ser Asn Glu Leu Gly Thr
 50 55 60
 Val Pro Leu Leu Lys Tyr Gly Ser Glu Glu Gln Arg Lys Arg Tyr Leu
 65 70 75 80
 Ser Glu Val Ala Ser Gly Lys Ala Leu Phe Gly Tyr Ala Leu Ser Glu
 85 90 95
 Ala Asp Ala Gly Ser Asp Pro Ala Ala Leu Lys Cys Arg Ala Asp Glu
 100 105 110
 Asp Gly Asp Ser Phe Val Leu Asn Gly Val Lys Ala Trp Val Thr Glu
 115 120 125
 Ala Gly Glu Ala Lys Tyr Leu Val Ile Phe Ala Val Thr Asp Pro Asp
 130 135 140
 Asp Pro Arg His Arg Ile Ser Ala Leu Met Val His Ala Asp Asp Pro
 145 150 155 160
 Gly Ile Ser Tyr Gly Ala Pro Glu His Lys Met Gly Ile Arg Gly Ser
 165 170 175
 Val Thr Arg Glu Val Val Phe Lys Asn Thr Arg Ile Pro Lys Glu Arg
 180 185 190
 Val Ile Gly Arg Arg Gly His Gly Leu Ser Val Ala Leu Gly Thr Leu
 195 200 205
 Asp Asn Ser Arg Val Ser Ile Ala Ala Gln Ala Val Gly Ile Ala Gln
 210 215 220
 Gly Ala Leu Asp Ile Ala Thr Asp Tyr Val Gln Lys Arg Lys Gln Phe
 225 230 235 240
 Gly Gln Pro Leu Ser Asn Phe Glu Gly Ile Gln Phe Met Leu Ala Asp
 245 250 255
 Met Ala Met Arg Leu Glu Ala Ala Arg Ala Leu Thr Tyr Ser Ala Ala
 260 265 270
 Asp Arg Ser Gly Arg Gln Thr Asp Asp Val Ser Tyr Phe Gly Ala Ala
 275 280 285
 Ala Lys Cys Phe Ala Ser Asp Thr Ala Met Ala Val Cys
 290 295 300

<210> 47
 <211> 379
 <212> DNA
 <213> Homo sapiens

<400> 47
 aagctttag agctagtcgg aagcggactg tcggtagccc aagctgctaa aagatgtggg
 60
 atgcatttta ccgctgcgta tgccgtagct acggaagctg ggtgccatat ccggttaagt
 120

cagtatgctc ggaaagtccg ccagacgcag ttaagagtgg aatacctgcg ccttcggctg
 180
 gcgagcctgc ctggtgggtga tgctggcgcg gcagtaggaa ttgatcgctg actgcgttta
 240
 gatttcgaaa aaggactcac caaatcccag ggctgcgcag aagagttcat acccgctcggc
 300
 gaagacgcca gcacgtataa cagacttatg aaagcgctgc gccaacgcca tgatgtcatc
 360
 aaatccggaa agcttgccc
 379

<210> 48
 <211> 106
 <212> PRT
 <213> Homo sapiens

<400> 48
 Met His Leu Thr Ala Ala Tyr Ala Val Ala Thr Glu Ala Gly Cys His
 1 5 10 15
 Ile Arg Leu Ser Gln Tyr Ala Arg Lys Val Arg Gln Thr Gln Leu Arg
 20 25 30
 Val Glu Tyr Leu Arg Leu Arg Leu Ala Ser Leu Pro Gly Gly Asp Ala
 35 40 45
 Gly Ala Ala Val Gly Ile Asp Arg Arg Leu Arg Leu Asp Phe Glu Lys
 50 55 60
 Gly Leu Thr Lys Ser Gln Gly Arg Arg Glu Glu Phe Ile Pro Val Gly
 65 70 75 80
 Glu Asp Ala Ser Thr Tyr Asn Arg Leu Met Lys Ala Leu Arg Gln Arg
 85 90 95
 His Asp Val Ile Lys Ser Gly Lys Leu Ala
 100 105

<210> 49
 <211> 309
 <212> DNA
 <213> Homo sapiens

<400> 49
 tgatcatgat gctggcatgg actattctgg tccctgttcc tctctcacct gctgaaggac
 60
 atccctctaa tttttgtgtc tccttctgta tcatcaaatt ttcctctct actgagtctc
 120
 ttgcatctcc ttggaagcat gctgtactat gtcccatcct taaagaactc cccttgctg
 180
 cacattaccc tctgccagct ggctcatttt tctgctcccc tttacagggg aactcttcaa
 240
 aaagttatct ccacctcctt ccattctatg ttctcttgaa cctgcagtac tgggtgctcc
 300
 ctctttttg
 309

<210> 50
 <211> 101
 <212> PRT

<213> Homo sapiens

<400> 50

```

Met Met Leu Ala Trp Thr Ile Leu Val Pro Val Pro Leu Ser Pro Ala
 1           5           10           15
Glu Gly His Pro Ser Asn Phe Cys Val Ser Phe Cys Ile Ile Lys Phe
      20           25           30
Ser Leu Ser Thr Glu Ser Leu Ala Ser Pro Trp Lys His Ala Val Leu
      35           40           45
Cys Pro Ile Leu Lys Glu Leu Pro Leu Ser Ala His Tyr Pro Leu Pro
      50           55           60
Ala Gly Ser Phe Phe Cys Ser Pro Leu Gln Gly Asn Ser Ser Lys Ser
65           70           75           80
Tyr Leu His Leu Leu Pro Ser His Val Leu Leu Asn Leu Gln Tyr Trp
      85           90           95
Val Leu Pro Pro Phe
      100

```

<210> 51

<211> 512

<212> DNA

<213> Homo sapiens

<400> 51

```

agatctttga agaattgcc cactgtcttc ctccctgctt ataatttcct tattccctag
60
gatgtgatcc ttgttcttgg ggcctcacat ggcagctgga tctctggcga ttgcatctga
120
gttccagaca ccaggatgga aaagaaaaga aggaggggca agaggaaccc ccagatgctc
180
cttaagagct actgcgtggc attcccactt gcatctcatt tgctcgatcg ctgtcactgt
240
gccctaacga gctgcaagga cactggggaa atgagtctgt cttgtacttc atgtgccctt
300
caaaatcttc tgttgctgag ggagaagagg ccagccggta ttgaggaaca actagcactt
360
tctgcttccg cgcccaggg ggacgtgggt gtgttgaatc cacaccgggg gtgaggacct
420
ctgaggctgg gctggatggg acatcaggtg ggcctctgt ttcatttatg tgacctccca
480
tcagggtcttc tggttggatc ctgctttcta ga
512

```

<210> 52

<211> 125

<212> PRT

<213> Homo sapiens

<400> 52

```

Met Glu Lys Lys Arg Arg Arg Gly Lys Arg Asn Pro Gln Met Leu Leu
 1           5           10           15
Lys Ser Tyr Cys Val Ala Phe Pro Leu Ala Ser His Leu Leu Asp Arg
      20           25           30
Cys His Cys Ala Leu Thr Ser Cys Lys Asp Thr Gly Glu Met Ser Leu

```

```

      35              40              45
Ser Cys Thr Ser Cys Ala Pro Gln Asn Leu Leu Leu Arg Glu Lys
      50              55              60
Arg Pro Ala Gly Ile Glu Gln Leu Ala Leu Ser Ala Ser Ala Ser
65              70              75              80
Gln Gly Asp Val Gly Val Leu Asn Pro His Arg Gly Cys Gly Pro Leu
      85              90              95
Arg Leu Gly Trp Met Gly His Gln Val Gly Pro Leu Phe His Leu Cys
      100             105             110
Asp Leu Pro Ser Gly Leu Leu Val Gly Ser Cys Phe Leu
      115             120             125

```

<210> 53
 <211> 474
 <212> DNA
 <213> Homo sapiens

```

<400> 53
accggtacac ctacgtcacc cgtaaaaacc gacgcaatac ccggatcgcc tcgtcctcaa
60
aaaattcgat cccgtcgtgc gtcgtcacat tgagttcaag gaggcccgct aatggccaaa
120
aagtccaaga ttgtcgccca gaagaaacgt gagaagctcg tagcccaata cgccgaaagg
180
cgcgccgaac tcaaggccat catgaagtgc ccaactgcct cattggacga acgcatggag
240
gcacgcgta agctgtctcg cctgccgcgc gattcatccc ccgtgcggtt acgtaaccgt
300
gaccaagtcg acgggcgctcc ccgcggctac gttggcaagg ccggtgtgtc ccgtatccgt
360
ttccgtgaga tggcccaccg cggcgaactc cccggaatcg cgaagtcaag ctggtgaagc
420
catggcagta ccgaagcgaa agaagtcccg ttcgaccacg cgatcataggc gggc
474

```

<210> 54
 <211> 101
 <212> PRT
 <213> Homo sapiens

```

<400> 54
Met Ala Lys Lys Ser Lys Ile Val Ala Gln Lys Lys Arg Glu Lys Leu
1              5              10              15
Val Ala Gln Tyr Ala Glu Arg Arg Ala Glu Leu Lys Ala Ile Met Lys
      20              25              30
Cys Pro Thr Ala Ser Leu Asp Glu Arg Met Glu Ala Ser Arg Lys Leu
      35              40              45
Ser Arg Leu Pro Arg Asp Ser Ser Pro Val Arg Leu Arg Asn Arg Asp
      50              55              60
Gln Val Asp Gly Arg Pro Arg Gly Tyr Val Gly Lys Ala Gly Val Ser
65              70              75              80
Arg Ile Arg Phe Arg Glu Met Ala His Arg Gly Glu Leu Pro Gly Ile
      85              90              95
Ala Lys Ser Ser Trp

```

100

<210> 55
 <211> 378
 <212> DNA
 <213> Homo sapiens

<400> 55
 ccatggccca ggacagccgg catatcggct acgactacgg tacaccggtg gcgccacagt
 60
 tcggcgcgagc caagcccgca gcgtgctgcc aggcgcaagc gacaaacacc ggcccgtggg
 120
 tgggtgttcga ccatgtgcgt tgcacccacg acacctttct gatcgacgtc tttctcaacc
 180
 agcccgatgc caccgcgag caggtcaatg ccgacaaccc gcactacgtc gggcgtttca
 240
 gccgcacggt catggggcctg gtggatgaca agggccggtt cattaccag ggcgtatcgc
 300
 gcgcggttgaa tgcggcgcg agcaccaagg cgctgaacct gggaccgagt gacgcggcgc
 360
 agttatcggg gaggcgta
 378

<210> 56
 <211> 125
 <212> PRT
 <213> Homo sapiens

<400> 56
 Met Ala Gln Asp Ser Arg His Ile Gly Tyr Asp Tyr Gly Thr Pro Val
 1 5 10 15
 Ala Pro Gln Phe Gly Ala Ala Lys Pro Ala Ala Cys Cys Gln Ala Gln
 20 25 30
 Ala Thr Asn Thr Gly Pro Trp Val Val Phe Asp His Val Arg Cys Thr
 35 40 45
 His Asp Thr Phe Leu Ile Asp Val Phe Leu Asn Gln Pro Asp Ala Thr
 50 55 60
 Ala Gln Gln Val Asn Ala Asp Asn Pro His Tyr Val Gly Arg Phe Ser
 65 70 75 80
 Arg Ile Gly Met Gly Leu Val Asp Asp Lys Gly Arg Cys Ile Thr Gln
 85 90 95
 Gly Val Ser Arg Ala Leu Asn Ala Ala Arg Ser Thr Lys Ala Leu Asn
 100 105 110
 Leu Gly Pro Ser Asp Ala Ala Gln Leu Ser Val Arg Arg
 115 120 125

<210> 57
 <211> 388
 <212> DNA
 <213> Homo sapiens

<400> 57
 agaccacccc gacacagatc aggagtcgtc atgtccagaa agaagaaggt cggcatcctc
 60

accgcaggcg gtgattgcc cgggctcaac gccgctatcc gcggatttgg caaggctgcc
 120
 atccgccagc acgacatgga gctcatcggt attcaggacg gctttcttgg attggcggga
 180
 aaccgcacca tctcccttgg cccgcgtgcc ctctcaggca tcttgacggt cggcgggacc
 240
 atcctgggaa ctagccgtga caaggccaat cacatgatta tcgacggcga ggaacgggat
 300
 atggtcccca ccacgctga gaattacgag aagctggggc ttgacgcttt ggtgactttg
 360
 ggtggcgggtg gcaccgcaa gaacgcgt
 388

<210> 58
 <211> 129
 <212> PRT
 <213> Homo sapiens

<400> 58
 Arg Pro Thr Arg His Arg Ser Gly Val Val Met Ser Arg Lys Lys Lys
 1 5 10 15
 Val Gly Ile Leu Thr Ala Gly Gly Asp Cys Pro Gly Leu Asn Ala Ala
 20 25 30
 Ile Arg Gly Phe Gly Lys Ala Ala Ile Arg Gln His Asp Met Glu Leu
 35 40 45
 Ile Gly Ile Gln Asp Gly Phe Leu Gly Leu Ala Gly Asn Arg Thr Ile
 50 55 60
 Ser Leu Gly Pro Arg Ala Leu Ser Gly Ile Leu Thr Val Gly Gly Thr
 65 70 75 80
 Ile Leu Gly Thr Ser Arg Asp Lys Val Asn His Met Ile Ile Asp Gly
 85 90 95
 Glu Glu Arg Asp Met Val Pro Thr Thr Val Glu Asn Tyr Glu Lys Leu
 100 105 110
 Gly Leu Asp Ala Leu Val Thr Leu Gly Gly Gly Thr Ala Lys Asn
 115 120 125
 Ala

<210> 59
 <211> 417
 <212> DNA
 <213> Homo sapiens

<400> 59
 ggtaccatcg gagctcgaca agaaatgggtt ggggtgaagtc gtggcttctg ctccacccag
 60
 tgccctcatg ggtcagccca cctgaatatc ttcatgcctg tgcatttctc ctgatgttca
 120
 cgtgtgccct gtgtttttac gcatctgtga tcgtgcaccc acgcgtctca gagaggagcc
 180
 cgtttgggaa tccggagaat gtgcgtggc ggaagagcgt cacacactgg aagcaaacct
 240
 cagaccgcgt ggacaagacc aaggatgaaa tggaacacga ggccttgggtg gaaggggaacc
 300

tggcaaccga ggcaagccta gtggttctgg acacactgga gatcatcgtg cagacgggtga
360
tgctttcaga agccccgggag agcgtcttgg gggcagtgcg gaaggttctg ctgtaca
417

<210> 60
<211> 101
<212> PRT
<213> Homo sapiens

<400> 60
Met Phe Thr Cys Ala Leu Cys Phe Tyr Ala Ser Val Ile Val His Pro
1 5 10 15
Arg Val Ser Glu Arg Ser Pro Phe Gly Asn Pro Glu Asn Val Arg Trp
20 25 30
Arg Lys Ser Val Thr His Trp Lys Gln Thr Ser Asp Arg Val Asp Lys
35 40 45
Thr Lys Asp Glu Met Glu His Glu Ala Leu Val Glu Gly Asn Leu Ala
50 55 60
Thr Glu Ala Ser Leu Val Val Leu Asp Thr Leu Glu Ile Ile Val Gln
65 70 75 80
Thr Val Met Leu Ser Glu Ala Arg Glu Ser Val Leu Gly Ala Val Leu
85 90 95
Lys Val Val Leu Tyr
100

<210> 61
<211> 304
<212> DNA
<213> Homo sapiens

<400> 61
agatcttcac agccttagac ttttttcatg ggtgccttac agttttggag gtccttatcc
60
gcacacatat ttgcaggctt ggagagagtg tgtgggggca tgtactttcg gtgggtcaag
120
tatgaagaag caggccttat aaacacatat tctgacctta acctgtactt cagaagagga
180
ccgctgactc accaaggagg cctgaaggac aaggcagcat ctctgtcttc acatgagtc
240
tcccctagac cgggccccatg gccaggcctg accacagagc tcccattgcc tttcctgcac
300
gcgt
304

<210> 62
<211> 92
<212> PRT
<213> Homo sapiens

<400> 62
Met Gly Ala Leu Gln Phe Trp Arg Ser Leu Ser Ala His Ile Phe Ala
1 5 10 15

Gly Leu Glu Arg Val Cys Gly Gly Met Tyr Phe Arg Trp Val Lys Tyr
 20 25 30
 Glu Glu Ala Gly Leu Ile Asn Thr Tyr Ser Asp Leu Asn Leu Tyr Phe
 35 40 45
 Arg Arg Gly Pro Leu Thr His Gln Gly Gly Leu Lys Asp Lys Ala Ala
 50 55 60
 Ser Leu Ser Ser His Glu Ser Ser Pro Arg Pro Gly Pro Trp Pro Gly
 65 70 75 80
 Leu Thr Thr Glu Leu Pro Leu Pro Phe Leu His Ala
 85 90

<210> 63

<211> 577

<212> DNA

<213> Homo sapiens

<400> 63

cgcgctcaagg ggggtctacac cgggacgatt aacgcctcgg tgggagtatt catcaccgcg
 60
 ctgacgggtgc tagctgggtg gctcacccta gccgggcgta tcagtgtcgg ggaactcgtc
 120
 accgtgggtcg ggctggccca aaccctcggc cctccgctgc gagcactggg cgtcgacacc
 180
 gcgacgatgt tggccaccgc ccacgcctcc ggggaccgat tctgtgagtt gcgtgatagc
 240
 ccggcagcct ggcagatcca ccccgacgac ggtgcccgca ccacaccggg tgatggcccg
 300
 gtggagttgc acatcccggc cagggatttc cagcttgacg tcgccggcgg caccatgtg
 360
 ggtatcatgg cgctcaatc ggtctgtgac gccttggccg aggcgataga ccacggctcc
 420
 gagaccgtct tgaatggggc tcccgcagc cgctcaacc ctgcccacg gcgtcgtctg
 480
 gtgctgggtg ctccccgctc ccccgaaactg ttcgacgata ctgcccgctg gaacatcgtg
 540
 cttgacagcc agacgactgt cgccaggctg aatgcat
 577

<210> 64

<211> 192

<212> PRT

<213> Homo sapiens

<400> 64

Arg Val Lys Gly Val Tyr Thr Gly Thr Ile Asn Ala Ser Val Gly Val
 1 5 10 15
 Phe Ile Thr Ala Leu Thr Val Leu Ala Gly Trp Leu Thr Leu Ala Gly
 20 25 30
 Arg Ile Ser Val Gly Glu Leu Val Thr Val Val Gly Leu Ala Gln Thr
 35 40 45
 Leu Gly Pro Pro Leu Arg Ala Leu Gly Val Asp Thr Ala Thr Met Leu
 50 55 60
 Ala Thr Ala His Ala Ser Gly Asp Arg Phe Cys Glu Leu Arg Asp Ser
 65 70 75 80

Pro Ala Ala Trp Gln Ile His Pro Asp Asp Gly Ala Arg Thr Thr Pro
 85 90 95
 Gly Asp Gly Pro Val Glu Leu His Ile Pro Val Arg Asp Phe Gln Leu
 100 105 110
 Asp Val Ala Gly Gly Thr His Val Gly Ile Met Ala Pro Gln Ser Val
 115 120 125
 Cys Asp Ala Leu Ala Glu Ala Ile Asp His Gly Ser Glu Thr Val Leu
 130 135 140
 Asn Gly Val Pro Ala Ser Arg Leu Asn Pro Ala Gln Arg Arg Arg Leu
 145 150 155 160
 Val Leu Val Ala Pro Arg Ser Pro Glu Leu Phe Asp Asp Thr Ala Arg
 165 170 175
 Ala Asn Ile Val Leu Asp Ser Gln Thr Thr Val Ala Arg Leu Asn Ala
 180 185 190

<210> 65

<211> 339

<212> DNA

<213> Homo sapiens

<400> 65

gtcgaccgcg ccttgggacg gctcgaaggg gccagcctgg accaggtagc ggaagaagtc
 60
 aagaaggccg ctttcaagat caccgcgcgc gggcaactag tgggcacccat ggcctccgag
 120
 cgccttggcg tacccttcgg catcatcgac ctttcgcttg cccctactgc cgaattggga
 180
 gattcgggggg cccacatcct tgagcatatg ggattggacc aagtaggcac gcacggcaca
 240
 actgctgctt tggctctgct taacgacgcc gtaaagaaag gcggcatgat ggctgcccc
 300
 cgcgtcgcg gtttgtctgg ctcccttcac ccgggctcc
 339

<210> 66

<211> 113

<212> PRT

<213> Homo sapiens

<400> 66

Val Asp Arg Ala Leu Gly Ser Leu Glu Gly Ala Ser Leu Asp Gln Val
 1 5 10 15
 Ala Glu Glu Val Lys Lys Ala Ala Phe Lys Ile Thr Arg Ala Gly Gln
 20 25 30
 Leu Val Gly Thr Met Ala Ser Glu Arg Leu Gly Val Pro Phe Gly Ile
 35 40 45
 Ile Asp Leu Ser Leu Ala Pro Thr Ala Glu Leu Gly Asp Ser Gly Ala
 50 55 60
 His Ile Leu Glu His Met Gly Leu Asp Gln Val Gly Thr His Gly Thr
 65 70 75 80
 Thr Ala Ala Leu Ala Leu Leu Asn Asp Ala Val Lys Lys Gly Gly Met
 85 90 95
 Met Ala Cys Pro Arg Val Gly Gly Leu Ser Gly Ser Phe Ile Pro Gly
 100 105 110

Ser

<210> 67

<211> 446

<212> DNA

<213> Homo sapiens

<400> 67

tgatcataaa ccacgcgtca ccgaggggat gtggcacacc tacctgcgcg tcgcagatgc
 60
 cgcacaggca cgggtcaggg gcgttcgcgg cgccagctgg cacaacttcg cgaccgggta
 120
 caaggggtcc ttcgacgcca acgagcttgc cgtaactcct gatactgaca ccgtcatcca
 180
 gggagtcggg cccgccctag cctcctcga ttcagcgtgg ggacgccaga tccacgtgga
 240
 gacaacaggg tgtcccagtg ccgtggtctg gaatccacgc tcctcgtcga cacatgccga
 300
 taaccgcgaca gccacggcat ggcgcgattt cgtatgcgtc gagaccgggg cctgcaagga
 360
 caatgcggtc attgttgccc cacacagcga cctcaccatg tccacacgga ttagcgtcga
 420
 aacgttgtga tcgctgcatg gatatt
 446

<210> 68

<211> 133

<212> PRT

<213> Homo sapiens

<400> 68

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Trp | His | Thr | Tyr | Leu | Arg | Val | Ala | Asp | Ala | Ala | Gln | Ala | Arg | Val |
| 1 | | | | | 5 | | | | 10 | | | | | 15 | |
| Arg | Gly | Val | Arg | Gly | Ala | Ser | Trp | His | Asn | Phe | Ala | Thr | Gly | Asp | Lys |
| | | | 20 | | | | | 25 | | | | | 30 | | |
| Gly | Ser | Phe | Asp | Ala | Asn | Glu | Leu | Ala | Val | Thr | Pro | Asp | Thr | Asp | Thr |
| | | 35 | | | | | 40 | | | | | 45 | | | |
| Val | Ile | Gln | Gly | Val | Gly | Pro | Ala | Leu | Ala | Leu | Leu | Asp | Ser | Ala | Trp |
| | | 50 | | | | 55 | | | | | 60 | | | | |
| Gly | Arg | Gln | Ile | His | Val | Glu | Thr | Thr | Gly | Cys | Pro | Ser | Ala | Val | Val |
| 65 | | | | | 70 | | | | 75 | | | | | 80 | |
| Trp | Asn | Pro | Arg | Ser | Ser | Ser | Thr | His | Ala | Asp | Asn | Pro | Thr | Ala | Gln |
| | | | | 85 | | | | 90 | | | | | 95 | | |
| Ala | Trp | Arg | Asp | Phe | Val | Cys | Val | Glu | Thr | Gly | Ala | Cys | Lys | Asp | Asn |
| | | | 100 | | | | | 105 | | | | | 110 | | |
| Ala | Val | Ile | Val | Ala | Pro | His | Ser | Asp | Leu | Thr | Met | Ser | Thr | Arg | Ile |
| | | 115 | | | | | 120 | | | | | | 125 | | |
| Ser | Val | Glu | Thr | Leu | | | | | | | | | | | |
| | | | | | | | | | | | | | | | |

<210> 69

<211> 552

<212> DNA

<213> Homo sapiens

<400> 69

nnaagggttaa ggagaaaagc aaggaccttg caaagagagc ctctgtgccg gagaggctgg
 60
 ccctcaagga ggagccaaaa gaagacccca gtggagcagc tgtgcccag atgccaaaaa
 120
 agtccctcaa gattgccagc ttcaccccca aaggggggaa gctcaacagt gccagaagg
 180
 agcnccatgg ccccttcctt cagtggaata ccaaaaccag gaatgaaaag catgcccggg
 240
 aaatccccaa gtgccccagc gccttccaag gaagggggagc ggagccggag tgggaagctg
 300
 agctcaggac tccccagca gaagccccag ctggacggca gacactccag ttcctcttcc
 360
 agcctggcgt cctcagaagg aaaaggccca ggagggacca ccctgaacca cagcatcagc
 420
 agccagactg tcagtgggtc tgtcgggacc acccagacca caggaagcaa tnnaccgtca
 480
 gtgttcagct acctcagccc cagcagcaat acaaccatcc caacactgcc acggttgac
 540
 ctttcctgta ca
 552

<210> 70

<211> 184

<212> PRT

<213> Homo sapiens

<400> 70

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Xaa | Arg | Val | Arg | Arg | Lys | Ala | Arg | Thr | Leu | Gln | Arg | Glu | Pro | Leu | Cys |
| 1 | | | | 5 | | | | | 10 | | | | | 15 | |
| Arg | Arg | Gly | Trp | Pro | Ser | Arg | Arg | Ser | Gln | Lys | Lys | Thr | Pro | Val | Glu |
| | | 20 | | | | | | 25 | | | | | 30 | | |
| Gln | Leu | Cys | Pro | Arg | Cys | Gln | Lys | Ser | Pro | Pro | Arg | Leu | Pro | Ala | Ser |
| | 35 | | | | | 40 | | | | | 45 | | | | |
| Ser | Pro | Lys | Gly | Gly | Ser | Ser | Thr | Val | Pro | Arg | Arg | Ser | Xaa | Met | Ala |
| | 50 | | | | 55 | | | | | 60 | | | | | |
| Pro | Ser | Leu | Ser | Gly | Ile | Pro | Lys | Pro | Gly | Met | Lys | Ser | Met | Pro | Gly |
| 65 | | | | 70 | | | | | 75 | | | | 80 | | |
| Lys | Ser | Pro | Ser | Ala | Pro | Ala | Pro | Ser | Lys | Glu | Gly | Glu | Arg | Ser | Arg |
| | | 85 | | | | | | 90 | | | | | 95 | | |
| Ser | Gly | Lys | Leu | Ser | Ser | Gly | Leu | Pro | Gln | Gln | Lys | Pro | Gln | Leu | Asp |
| | 100 | | | | | | 105 | | | | | 110 | | | |
| Gly | Arg | His | Ser | Ser | Ser | Ser | Ser | Leu | Ala | Ser | Ser | Glu | Gly | Lys | |
| | 115 | | | | | 120 | | | | | 125 | | | | |
| Gly | Pro | Gly | Gly | Thr | Thr | Leu | Asn | His | Ser | Ile | Ser | Ser | Gln | Thr | Val |
| | 130 | | | | | 135 | | | | | 140 | | | | |
| Ser | Gly | Ser | Val | Gly | Thr | Thr | Gln | Thr | Thr | Gly | Ser | Asn | Xaa | Pro | Ser |
| 145 | | | | 150 | | | | | 155 | | | | | 160 | |
| Val | Phe | Ser | Tyr | Leu | Ser | Pro | Ser | Ser | Asn | Thr | Thr | Ile | Pro | Thr | Leu |
| | | | 165 | | | | | 170 | | | | | 175 | | |
| Pro | Arg | Leu | His | Leu | Ser | Cys | Thr | | | | | | | | |
| | | | 180 | | | | | | | | | | | | |

<210> 71
 <211> 316
 <212> DNA
 <213> Homo sapiens

<400> 71
 cgcgttgaaa tggcgttcga acttaaacgt ttacatattg actccgtgcc attaaacatt
 60
 ttgaatcctg ttaaaggagac tccatttgaa agcaacgaag ctttacgtcc tttaaataac
 120
 ttacgtacct tcgccgtatt ccgtttcatc ttgccaaacg cattgatacg aactgcaggt
 180
 ggccgcgaag taaatctacg agacttacaa gcttatgctc taaaagggtg cctaaacggt
 240
 atcatggttg gtggctactt aactactggc ggtcgttcac ctcaagacga tctccaaatg
 300
 attcaagact tggagt
 316

<210> 72
 <211> 105
 <212> PRT
 <213> Homo sapiens

<400> 72
 Arg Val Glu Met Ala Phe Glu Leu Lys Arg Leu His Ile Asp Ser Val
 1 5 10 15
 Pro Leu Asn Ile Leu Asn Pro Val Lys Gly Thr Pro Phe Glu Ser Asn
 20 25 30
 Glu Ala Leu Arg Pro Leu Asn Ile Leu Arg Thr Phe Ala Val Phe Arg
 35 40 45
 Phe Ile Leu Pro Asn Ala Leu Ile Arg Thr Ala Gly Gly Arg Glu Val
 50 55 60
 Asn Leu Arg Asp Leu Gln Ala Tyr Ala Leu Lys Gly Gly Leu Asn Gly
 65 70 75 80
 Ile Met Val Gly Gly Tyr Leu Thr Thr Gly Gly Arg Ser Pro Gln Asp
 85 90 95
 Asp Leu Gln Met Ile Gln Asp Leu Glu
 100 105

<210> 73
 <211> 384
 <212> DNA
 <213> Homo sapiens

<400> 73
 nntaccggca agatcctggc cgaagggtgac gtcgaggttt ctgaggctat cgactttgct
 60
 gcttggtatg tcgaccgagc cgaggagctc gagggcgctc acggtgcca gtttgtgccg
 120
 ccacgagtga ccgtcgtcac cccgccgtgg aacttcgccc tgtctattac cgccggatcc
 180

acccttgccg ctctggccgc cggatcgctca gtactactca agcccgtcc acaggcccg
 240
 cactgtgctg ccgtcatctc tgaatgcctg tgggaggctg ggatcccgcg ggacgttctg
 300
 cagctcgtcg atgttgagga aaatgaggct ggtaaaccac tggtagacca ccccgaggtc
 360
 gatcgggtca tcctcacggg aggt
 384

<210> 74.
 <211> 128
 <212> PRT
 <213> Homo sapiens

<400> 74
 Xaa Thr Gly Lys Ile Leu Ala Glu Gly Asp Val Glu Val Ser Glu Ala
 1 5 10 15
 Ile Asp Phe Ala Ala Trp Tyr Val Asp Arg Ala Glu Glu Leu Glu Gly
 20 25 30
 Val Asp Gly Ala Gln Phe Val Pro Pro Arg Val Thr Val Val Thr Pro
 35 40 45
 Pro Trp Asn Phe Ala Leu Ser Ile Thr Ala Gly Ser Thr Leu Ala Ala
 50 55 60
 Leu Ala Ala Gly Ser Ser Val Leu Leu Lys Pro Ala Pro Gln Ala Arg
 65 70 75 80
 His Cys Ala Ala Val Ile Ser Glu Cys Leu Trp Glu Ala Gly Ile Pro
 85 90 95
 Arg Asp Val Leu Gln Leu Val Asp Val Glu Glu Asn Glu Ala Gly Lys
 100 105 110
 His Leu Val Ser His Pro Glu Val Asp Arg Val Ile Leu Thr Gly Gly
 115 120 125

<210> 75
 <211> 405
 <212> DNA
 <213> Homo sapiens

<400> 75
 gaattcgtct cggaatacac gctggaaaat tcggccgaga tgtccggggt gcgctcanac
 60
 cgcattgagg cgctggccga gctctatgcc gatcccaaga ccagggtggt gagcttctgg
 120
 accatgggct tcaaccagca caccgcggc gtctggtgca acaatctcgt ctacaacatc
 180
 cacctgctga ccggaaaaat ctgcagccc ggcaacagcc cggttctcgt gaccgggcag
 240
 ccacggcct ggggcacggc gcgcgaggtc ggtaccttct cgcacgcct gcccgccgac
 300
 atggtggtca ccagcaaggc gcaccgcgac atcgccgaga agatctggca gctgccggaa
 360
 ggaccagtcc ccgacaagcc cggctaccac gccgtgctgc agagc
 405

<210> 76

<211> 135
 <212> PRT
 <213> Homo sapiens

<400> 76
 Glu Phe Val Ser Glu Tyr Thr Leu Glu Asn Ser Ala Glu Met Ser Gly
 1 5 10 15
 Val Arg Ser Xaa Arg Ile Glu Ala Leu Ala Glu Leu Tyr Ala Asp Pro
 20 25 30
 Lys Thr Arg Val Val Ser Phe Trp Thr Met Gly Phe Asn Gln His Thr
 35 40 45
 Arg Gly Val Trp Cys Asn Asn Leu Val Tyr Asn Ile His Leu Leu Thr
 50 55 60
 Gly Lys Ile Ser Thr Pro Gly Asn Ser Pro Phe Ser Leu Thr Gly Gln
 65 70 75 80
 Pro Ser Ala Cys Gly Thr Ala Arg Glu Val Gly Thr Phe Ser His Arg
 85 90 95
 Leu Pro Ala Asp Met Val Val Thr Ser Lys Ala His Arg Asp Ile Ala
 100 105 110
 Glu Lys Ile Trp Gln Leu Pro Glu Gly Pro Val Pro Asp Lys Pro Gly
 115 120 125
 Tyr His Ala Val Leu Gln Ser
 130 135

<210> 77
 <211> 5816
 <212> DNA
 <213> Homo sapiens

<400> 77
 gagcggcgcc ctgctctggc cggtgctccc gctcctgctc ctgctgctgt cggcgcggga
 60
 cggcgtgcmc gccgcgcagc ctcaggcccc gggttacttg attgcagctc cctctgtttt
 120
 tcgcgcgggc gtggaggaag tcatacagct gaccatcttt aactctccaa gggaagtcac
 180
 ggtccaggct cagctggtgg cccagggtga gccggtggtg cagagccagg gagccatcct
 240
 ggataaagg acaatcaaac tcaagggtgc cacgggcctc cggggccaag cgcttctgaa
 300
 agtgtggggc cgcggctggc aggcggagga ggggcccctc ttccacaacc agacctcggt
 360
 gaccgtggac ggccggggcg cttctgtatt catccagacg gacaagcctg tgtacagacc
 420
 ccagcaccga gtgctcataa gcatcttcac cgtctctcca aatctgaggc ctgtcaacga
 480
 gaagctggaa gcctacatcc tggacccccg aggtctctcg atgatagagt ggagacactt
 540
 aaagccgttc tgetgcggca tcaccaacat gagcttcccc ttgtccgacc agcctgtggt
 600
 gggagaatgg ttcatttttg ttgaaatgca aggccacgcg tacaacaagt cttttgaagt
 660
 tcagaagtat gtgttgccca agtttgagct tctgattgac ccgccccggt atatccaaga
 720

cctggacgcc tgtgagacag gcactgtgcg ggccaggtat acctttggga aacctgtggc
780
tggtgcctta atgatcaaca tgactgttaa tgggtgtaggg tactacagcc acgaggtggg
840
acgccctgtc ctcaagaaca ccaagatcct cggctcccgg gacttcgaca tctgctgag
900
ggacatgatc ccagcggacg tccctgagca cttccggggc agggtcagca tctgggccat
960
ggtgaccagt gtggacggga gccagcaggt cgcgttcgat gactccaccc ccgtgcagag
1020
gcagctggtg gacatccggt actccaagga cagcaggaag cagttcaagc cgggcctggc
1080
ctacgtgggg aaggtggagc tatcctaccc cgatggcagc ccagctgagg gggtgacggt
1140
ccagattaag gcagagctga caccaaagga taacatctac accagtgaag ttgtgtccca
1200
gcgtggacta gtgggggtttg aaatcccctc catcccacg tcagcccagc acgtgtggct
1260
ggagaccaag gtgatggcac tgaacgggaa gcccgtaggg gctcagtacc tgcccagcta
1320
cctctcccctc ggcagctggt actccccag ccagtgtctac ctgcagctgc agccaccctc
1380
ccaccactg caggttgggg aagaagccta tttttctgtg aagtccacat gtccctgcaa
1440
ctttaccctg tactacgagg tggctgcacg gggcaatatt gtgctatcgg gccagcagcc
1500
tgccacacac acccagcagc gaagcaagcg ggcggcccct gccctggaga aaccgattcg
1560
tttaacacac ctttctgaga cagagcccc accagcccca gaagctgagg tcgacgtgtg
1620
tgtgacctct cttcatctgg ccgtgacccc cagcatggtc ccccttggtc gcctgctggt
1680
ctttctacgtc agggagaatg gagaaggggt cgcgcacagc cttcagtttg cagtcgagac
1740
cttcttcgaa aaccaggttt cagtgcgta ttcagcaaat gagacccaac ctggggaggt
1800
tgtgcacctg cggatcaggg ctgcaagggg cagctgtgtg tgcgtcgccg cagttgataa
1860
gagtgtctac ctgctcaggt ctgggttccg gctgactcct gccaggttt tccaggaact
1920
ggaagattat gatgtttctg attcctttgg cgtgtccagg gaggatggtc ctttttggtg
1980
ggctgggctg acggcacaac gacgccggcg ctcctctgtc ttcccgtggc cttggggcat
2040
caccaaggac tctgggtttg ctttcaccga aacgggactg gtggtgatga ccgaccgagt
2100
gagcctgaac caccggcagg acggtggcct ctacaccgat gaggtgtgcc ccgctttcca
2160
gccccacaca gggagcctgg tggcagtggc tccttccagg ccccccca gaacagagaa
2220
gagaaaaagg actttcttcc ccgaaacatg gatttggcat tgtctcaaca tcagtgaccc
2280
atctggtgag gggacactca gtgtgaaggt cccggactcc atcaccagct gggtaggtga
2340

ggcgtggcc ctgtccacct ctcagggtt aggcacgcc gagccctccc tgctgaagac
2400
cttcaagccc ttcttcgtgg acttcatgct ccccgctctc atcatccgtg gggagcaggt
2460
caagatcccc ctcagtgtct acaactacat gggcacctgc gctgagggtg acatgaagct
2520
ctcggttccc aagggcaccc agtttgttgg gcatcctggc aaacgccatg tgaccaagaa
2580
gatgtgtgtg gccccgggg aggctgagcc catctgggtc gttctgtcct tcagcgacct
2640
gggactcaac aacatcacgg ccaaagccct tgcttacgga gacacaaatt gctgccggga
2700
tgggaggtcc agcaaacc ctgaggagaa tcacgccgac aggagggtcc ccatcggggg
2760
ggatcacgtc aggcgcagtg tgatggttga ggcggaagga gtccccggg cgtacaccta
2820
cagcgcatcc ttctgtccca gtgagagagt ccacatctcc accccaaca agtatgagtt
2880
ccagtatgtg cagcgccac tgcgcctcac ccgctttgat gtggctgtgc gagctcacia
2940
tgatgccgt gtggccttgt cttctgggcc ccaggacaca gcaggcatga tcgagatcgt
3000
cctggggggg catcagaaca ccaggtcagt gatctccacc agcaagatgg gagagcccgt
3060
ggccagtga cacacggcca agatcctctc ctgggatgaa ttcagaacat tctggatcag
3120
ctggcgtggt ggccttatcc aggttgccca tggccagag ccatccaatg agtctgtcat
3180
tgtggcctgg accctccga ggcaccaga ggtccagttc attggctttt ccaccggctg
3240
gggctccatg ggtgaattcc gaatctggag gaagatggag gtggacgaga gctacagcga
3300
ggccttcacc tgggggtccc acacgggcc atccctgggt ctgagcgagc caccgcctcc
3360
atcatcgggg acgtcatggg gccaacctg aaccacctca acaacctcct gcggctgccg
3420
tttggtgtg gagagcagaa catgatccac ttgaccca acgtctttgt cttgaagtat
3480
cttcagaaaa ccagcagct cagccctgag gtggagagag agaccaccga ctacctagta
3540
caaggctacc agcgccagct gacctacaag cgccaggatg gctcctacag cgcgtttggg
3600
gagcgggacg catcggggag catgtggctc acagcctttg tctgaagtc cttgcacag
3660
gctcgagct ttatcttcgt ggacccccg gagctggctg ccgccaagag ctggatcatc
3720
cagcagcagc aggcgatgg ctccttcctg gccgtgggca gggctcctgaa caaggacatc
3780
cagggtggga tccacggcat tgtcccgtg acagcctacg tgggtggtgc tctcctggaa
3840
acaggcacag cctcagagga ggagagaggc tccactgaca aagcgaggca cttcctggag
3900
tctgctgcgc ccctggccat ggaccttat agctgtgccc tgactaccta cgcgctgacc
3960

ctgctccgca gcccggcagc ccttgaggca ctgcgcaagc tccgtagcct ggccatcatg
4020
cgagatgggg tcacccactg gagcctgtca aattcctggg acgtggacaa gggcacattc
4080
ttgagcttca gtgacagggg ctctcagtca gtgggtctcg ccgaggtgga aatgacagcc
4140
tacgcccttc tgacctacac tctgctgggt gacgtggctg ccgccctgcc tgtggtgaag
4200
tggctgtccc agcagcgaaa tgcaactggg ggcttctcct ccactcagga cacctgcgtg
4260
gctctgcagg ccttggtgta atatgccatc ttgtcctatg ctggaggcat caacctcact
4320
gtctccctgg cctccaccaa cctggactac caggaaacct tcgagctgca caggaccaac
4380
cagaaggttc tgcagacagc agcgatcccc agcctcccca cggggctggt tgtgagtgc
4440
aagggggacg gctgctgcct gatgcagatt gatgtcacct acaatgtgcc tgacccggtg
4500
gccaagccag ctttccagct gctcgtaagc ctccaggagc ctgaggccca gggacgccc
4560
ccccccatgc ctgcctccgc agctgagggg tcccaggag actggccccc agctgacgat
4620
gatgaccag cgccgatca gcatcaccag gaatacaagg tgatgctgga ggtgtgcacc
4680
aggtggctgc atgcagggc ttccaatatg gctgtcctgg aggtgcccct gctgtcaggc
4740
ttccggcag acatcgagag cctggagcag ctgctccttg acaagcacat ggggatgaag
4800
aggtatgaag tggctggacg ccgagtgtc ttctactttg atgagatccc cagccggtgc
4860
ctgacgtgcg tgcggttccg tgctctccgg gagtgcgtgg tgggcaggac gtcggcgctg
4920
ccagtctccg tgtacgacta ctacgaacct gccttcgagg ccactcgctt ctacaacgtc
4980
agcacgcaca gccactcgc ccgggaactg tgcgccggac ccgctgcaa cgaagtggag
5040
cgcgcccctg ccgggggccc gggctggttc cccggcgagt cgggccctgc cgtggcccct
5100
gaggaggggg cggcgatcgc gcgatgcggc tgcgaccacg actgcggcgc ccaggggaac
5160
ccggtgtgcg gctccgacgg ggtggtctac gccagcgct gccgctgcg ggaggccgcc
5220
tgccgccagg ccgcgccct ggagcccgcg cctcccagct gctgcgccct cgagcagcgg
5280
ctgccggcct cgtcgtcctc cacctacggg gatgacctgg cttctgtggc cccggggcct
5340
ttacagcagg acgtgaagct gaatggagcc ggccttgagg tggaggactc agacctgag
5400
cctgaagggg aggcggagga cagggtcaca gccgggcctc ggcctcctgt gagcagcggg
5460
aacctggaaa gcagcaccca gagcgccagc ccgttcacaa gatggggcca gactccggcc
5520
cctcagagac atagtggccg ggtggtgggg gccacaggc cagggttct gagccctgtc
5580

ttcgtctaca gccacgctt tcagagtggg ggggaggagg gtttatggat gtcaaacacc
 5640
 tgcaccttga gataatccta caaccacatg cagttgtggg accgcagttt ggtcctgggg
 5700
 accattcata cccacacacc cagcttgtgc ctgtgggttaa catctcagaa aactctggta
 5760
 aatgatcact ccaggatatt gacacgaata caggttactg atcttactca catgtt
 5816

<210> 78

<211> 799

<212> PRT

<213> Homo sapiens

<400> 78

| | | | | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Ile | Pro | Asn | Leu | Glu | Glu | Asp | Gly | Gly | Gly | Arg | Glu | Leu | Gln | Arg | Gly | 1 | 5 | 10 | 15 |
| Leu | His | Leu | Gly | Val | Pro | His | Gly | Ala | Ile | Pro | Gly | Ser | Glu | Arg | Ala | 20 | 25 | 30 | |
| Thr | Ala | Ser | Ile | Ile | Gly | Asp | Val | Met | Gly | Pro | Thr | Leu | Asn | His | Leu | 35 | 40 | 45 | |
| Asn | Asn | Leu | Leu | Arg | Leu | Pro | Phe | Gly | Cys | Gly | Glu | Gln | Asn | Met | Ile | 50 | 55 | 60 | |
| His | Phe | Ala | Pro | Asn | Val | Phe | Val | Leu | Lys | Tyr | Leu | Gln | Lys | Thr | Gln | 65 | 70 | 75 | 80 |
| Gln | Leu | Ser | Pro | Glu | Val | Glu | Arg | Glu | Thr | Thr | Asp | Tyr | Leu | Val | Gln | 85 | 90 | 95 | |
| Gly | Tyr | Gln | Arg | Gln | Leu | Thr | Tyr | Lys | Arg | Gln | Asp | Gly | Ser | Tyr | Ser | 100 | 105 | 110 | |
| Ala | Phe | Gly | Glu | Arg | Asp | Ala | Ser | Gly | Ser | Met | Trp | Leu | Thr | Ala | Phe | 115 | 120 | 125 | |
| Val | Leu | Lys | Ser | Phe | Ala | Gln | Ala | Arg | Ser | Phe | Ile | Phe | Val | Asp | Pro | 130 | 135 | 140 | |
| Arg | Glu | Leu | Ala | Ala | Ala | Lys | Ser | Trp | Ile | Ile | Gln | Gln | Gln | Gln | Ala | 145 | 150 | 155 | 160 |
| Asp | Gly | Ser | Phe | Leu | Ala | Val | Gly | Arg | Val | Leu | Asn | Lys | Asp | Ile | Gln | 165 | 170 | 175 | |
| Gly | Gly | Ile | His | Gly | Ile | Val | Pro | Leu | Thr | Ala | Tyr | Val | Val | Val | Ala | 180 | 185 | 190 | |
| Leu | Leu | Glu | Thr | Gly | Thr | Ala | Ser | Glu | Glu | Glu | Arg | Gly | Ser | Thr | Asp | 195 | 200 | 205 | |
| Lys | Ala | Arg | His | Phe | Leu | Glu | Ser | Ala | Ala | Pro | Leu | Ala | Met | Asp | Pro | 210 | 215 | 220 | |
| Tyr | Ser | Cys | Ala | Leu | Thr | Thr | Tyr | Ala | Leu | Thr | Leu | Leu | Arg | Ser | Pro | 225 | 230 | 235 | 240 |
| Ala | Ala | Pro | Glu | Ala | Leu | Arg | Lys | Leu | Arg | Ser | Leu | Ala | Ile | Met | Arg | 245 | 250 | 255 | |
| Asp | Gly | Val | Thr | His | Trp | Ser | Leu | Ser | Asn | Ser | Trp | Asp | Val | Asp | Lys | 260 | 265 | 270 | |
| Gly | Thr | Phe | Leu | Ser | Phe | Ser | Asp | Arg | Val | Ser | Gln | Ser | Val | Val | Ser | 275 | 280 | 285 | |
| Ala | Glu | Val | Glu | Met | Thr | Ala | Tyr | Ala | Leu | Leu | Thr | Tyr | Thr | Leu | Leu | 290 | 295 | 300 | |
| Gly | Asp | Val | Ala | Ala | Ala | Leu | Pro | Val | Val | Lys | Trp | Leu | Ser | Gln | Gln | | | | |

| | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 305 | | | | | 310 | | | | | 315 | | | | 320 |
| Arg | Asn | Ala | Leu | Gly | Gly | Phe | Ser | Ser | Thr | Gln | Asp | Thr | Cys | Val |
| | | | | 325 | | | | | 330 | | | | | 335 |
| Leu | Gln | Ala | Leu | Ala | Glu | Tyr | Ala | Ile | Leu | Ser | Tyr | Ala | Gly | Gly |
| | | | 340 | | | | | | 345 | | | | 350 | |
| Asn | Leu | Thr | Val | Ser | Leu | Ala | Ser | Thr | Asn | Leu | Asp | Tyr | Gln | Glu |
| | 355 | | | | | | 360 | | | | 365 | | | |
| Phe | Glu | Leu | His | Arg | Thr | Asn | Gln | Lys | Val | Leu | Gln | Thr | Ala | Ile |
| | 370 | | | | | | 375 | | | | 380 | | | |
| Pro | Ser | Leu | Pro | Thr | Gly | Leu | Phe | Val | Ser | Ala | Lys | Gly | Asp | Gly |
| 385 | | | | | 390 | | | | | 395 | | | | 400 |
| Cys | Leu | Met | Gln | Ile | Asp | Val | Thr | Tyr | Asn | Val | Pro | Asp | Pro | Val |
| | | | 405 | | | | | | 410 | | | | | 415 |
| Lys | Pro | Ala | Phe | Gln | Leu | Leu | Val | Ser | Leu | Gln | Glu | Pro | Glu | Ala |
| | | 420 | | | | | | 425 | | | | | 430 | |
| Gly | Arg | Pro | Pro | Pro | Met | Pro | Ala | Ser | Ala | Ala | Glu | Gly | Ser | Arg |
| | 435 | | | | | | 440 | | | | | 445 | | |
| Asp | Trp | Pro | Pro | Ala | Asp | Asp | Asp | Asp | Pro | Ala | Ala | Asp | Gln | His |
| | 450 | | | | | 455 | | | | | 460 | | | |
| Gln | Glu | Tyr | Lys | Val | Met | Leu | Glu | Val | Cys | Thr | Arg | Trp | Leu | His |
| 465 | | | | | 470 | | | | | 475 | | | | 480 |
| Gly | Ser | Ser | Asn | Met | Ala | Val | Leu | Glu | Val | Pro | Leu | Leu | Ser | Gly |
| | | | 485 | | | | | | 490 | | | | | 495 |
| Arg | Ala | Asp | Ile | Glu | Ser | Leu | Glu | Gln | Leu | Leu | Leu | Asp | Lys | His |
| | | 500 | | | | | | 505 | | | | | 510 | |
| Gly | Met | Lys | Arg | Tyr | Glu | Val | Ala | Gly | Arg | Arg | Val | Leu | Phe | Tyr |
| | 515 | | | | | | 520 | | | | | 525 | | |
| Asp | Glu | Ile | Pro | Ser | Arg | Cys | Leu | Thr | Cys | Val | Arg | Phe | Arg | Ala |
| | 530 | | | | | 535 | | | | | 540 | | | |
| Arg | Glu | Cys | Val | Val | Gly | Arg | Thr | Ser | Ala | Leu | Pro | Val | Ser | Val |
| 545 | | | | | 550 | | | | | 555 | | | | 560 |
| Asp | Tyr | Tyr | Glu | Pro | Ala | Phe | Glu | Ala | Thr | Arg | Phe | Tyr | Asn | Val |
| | | | 565 | | | | | | 570 | | | | | 575 |
| Thr | His | Ser | Pro | Leu | Ala | Arg | Glu | Leu | Cys | Ala | Gly | Pro | Ala | Cys |
| | | | 580 | | | | | 585 | | | | | 590 | |
| Glu | Val | Glu | Arg | Ala | Pro | Ala | Arg | Gly | Pro | Gly | Trp | Phe | Pro | Gly |
| | 595 | | | | | | 600 | | | | 605 | | | |
| Ser | Gly | Pro | Ala | Val | Ala | Pro | Glu | Glu | Gly | Ala | Ala | Ile | Ala | Arg |
| | 610 | | | | | 615 | | | | | 620 | | | |
| Gly | Cys | Asp | His | Asp | Cys | Gly | Ala | Gln | Gly | Asn | Pro | Val | Cys | Gly |
| 625 | | | | | 630 | | | | | 635 | | | | 640 |
| Asp | Gly | Val | Val | Tyr | Ala | Ser | Ala | Cys | Arg | Leu | Arg | Glu | Ala | Cys |
| | | | 645 | | | | | | 650 | | | | | 655 |
| Arg | Gln | Ala | Ala | Pro | Leu | Glu | Pro | Ala | Pro | Pro | Ser | Cys | Cys | Ala |
| | | 660 | | | | | | 665 | | | | | 670 | |
| Glu | Gln | Arg | Leu | Pro | Ala | Ser | Ser | Ser | Ser | Thr | Tyr | Gly | Asp | Asp |
| | 675 | | | | | | 680 | | | | | 685 | | |
| Ala | Ser | Val | Ala | Pro | Gly | Pro | Leu | Gln | Gln | Asp | Val | Lys | Leu | Asn |
| | 690 | | | | | 695 | | | | | 700 | | | |
| Ala | Gly | Leu | Glu | Val | Glu | Asp | Ser | Asp | Pro | Glu | Pro | Glu | Gly | Glu |
| 705 | | | | | 710 | | | | | 715 | | | | 720 |
| Glu | Asp | Arg | Val | Thr | Ala | Gly | Pro | Arg | Pro | Pro | Val | Ser | Ser | Gly |
| | | | 725 | | | | | | 730 | | | | | 735 |
| Leu | Glu | Ser | Ser | Thr | Gln | Ser | Ala | Ser | Pro | Phe | His | Arg | Trp | Gly |

<400> 81

nngettttct cncctcgca cctgatctgc tgcagacgtg cgggaggggc attaaaagat
 60
 tgacttaaat ttgataccag aaatgtcagt gtggctttgc ttgaacagcc gcgtgcgggg
 120
 cctctatgag tggaatccag tctcatggcc ccccccattg ctctgttac cctggaggag
 180
 gctactccat gaggctggct ccggtggcac tgcctcgggg ctgcatgtac atgtgtgtgc
 240
 atgnntgtgt gcatgtgcgt gtgcacgtgt nngtgtgtgc ccgtgtgcat gtgcccgtgt
 300
 gcgtgtctgt gctgtgtgtg tgcgtgcatg tgtgcgtgtc tgtgcgtgnc tgtgtgtgtg
 360
 cgtgcatgtg tgtgcacgtg tgtgcatgtg cacgtgtgag caccctgtgt catgtccgtg
 420
 tgcacgcgt
 429

<210> 82
 <211> 79
 <212> PRT
 <213> Homo sapiens

<400> 82
 Gly Trp Leu Arg Trp His Cys Leu Gly Ala Ala Cys Thr Cys Val Cys
 1 5 10 15
 Met Xaa Val Cys Met Cys Val Cys Thr Cys Xaa Cys Val Pro Val Cys
 20 25 30
 Met Cys Pro Cys Ala Cys Leu Cys Cys Val Cys Ala Cys Met Cys Ala
 35 40 45
 Cys Leu Cys Val Xaa Val Cys Val Arg Ala Cys Val Cys Thr Cys Val
 50 55 60
 His Val His Val Cys Ala Pro Val Cys Met Ser Val Cys Thr Arg
 65 70 75

<210> 83
 <211> 411
 <212> DNA
 <213> Homo sapiens

<400> 83
 gtgcacttct tcggacgagg gtccttcggc gcccgcgata acgagagggg gcggagattt
 60
 cgtgcgctct ttaaaaagcg caaaggcttt tataagttcg atgtgctttt tttcaggccc
 120
 ggaaagacgc gatccgtaga taaaataggc ccgctttatc gcaaattggct ttatggttac
 180
 cgtttccgca tcgatttgca gctgcggaaa aaaaggcctg tggtcgatgc cgttgtggat
 240
 gatttcgatt ttcgctttgt cgatgccaaag ccctatcaaa tcgtttcgca aaaactcgct
 300
 cgcggcgacg atcttttgca cctttgtaag cccccgcttt atctgcatgc gcatcatgcg
 360
 gtcgctcttt tctttgagca tgacggaaac gatggaattg acgacggcga c
 411

<210> 84
 <211> 127
 <212> PRT
 <213> Homo sapiens

<400> 84
 Met Leu Lys Glu Lys Ser Asp Arg Met Met Arg Met Gln Ile Lys Arg
 1 5 10 15
 Gly Leu Thr Lys Val Gln Lys Ile Val Ala Ala Ser Glu Phe Leu Arg
 20 25 30
 Asn Asp Leu Ile Gly Leu Gly Ile Asp Lys Ala Lys Ile Glu Ile Ile
 35 40 45
 His Asn Gly Ile Asp His Arg Pro Phe Phe Pro Gln Leu Gln Ile Asp
 50 55 60
 Ala Glu Thr Val Thr Ile Lys Pro Phe Ala Ile Lys Arg Pro Tyr Phe
 65 70 75 80
 Ile Tyr Gly Ser Arg Leu Ser Gly Pro Glu Lys Lys His Ile Glu Leu
 85 90 95
 Ile Lys Ala Phe Ala Leu Phe Lys Glu Arg Thr Lys Ser Pro His Pro
 100 105 110
 Leu Val Ile Ala Gly Ala Glu Gly Pro Ser Ser Glu Glu Val His
 115 120 125

<210> 85
 <211> 333
 <212> DNA
 <213> Homo sapiens

<400> 85
 nnacgcgtgc cgcgcaaggg aaccatgttc gtgtcgggtgc gcgaaaccga caaggcgcgc
 60
 atcctcgacg cggtgaaact gctgagttcg ctccggttca aggtgatcgc cacctcgggc
 120
 acccagcgtt tccgtgtgga gaacggagta ccggcggaaa agatcaacaa ggtgctggaa
 180
 ggccgcccgc acattgttga cgcaattacc aacggcggagg tgcaactcgt tttcaatacc
 240
 accgaggggc cacaggcgtt ggctgacagc cgctcgttgc gacgcgctgc cctcttgcac
 300
 aaagtgccat attacaccac tctttcaggt gca
 333

<210> 86
 <211> 111
 <212> PRT
 <213> Homo sapiens

<400> 86
 Xaa Arg Val Pro Arg Lys Gly Thr Met Phe Val Ser Val Arg Glu Thr
 1 5 10 15
 Asp Lys Ala Arg Ile Leu Asp Ala Val Lys Leu Leu Ser Ser Leu Gly
 20 25 30
 Phe Lys Val Ile Ala Thr Ser Gly Thr Gln Arg Phe Leu Val Glu Asn

ATTORNEY DOCKET NO.: 15966-543

```

      35          40          45
Gly Val Pro Ala Glu Lys Ile Asn Lys Val Leu Glu Gly Arg Pro His
      50          55          60
Ile Val Asp Ala Ile Thr Asn Gly Glu Val Gln Leu Val Phe Asn Thr
65          70          75          80
Thr Glu Gly Pro Gln Ala Leu Ala Asp Ser Arg Ser Leu Arg Arg Ala
      85          90          95
Ala Leu Leu His Lys Val Pro Tyr Tyr Thr Thr Leu Ser Gly Ala
      100          105          110

```

<210> 87
 <211> 355
 <212> DNA
 <213> Homo sapiens

```

<400> 87
acgcgtgagg aaatgggggc cgcaggcctg gatcgcaagg tatggcagtg cccggtcgtc
60
ctcctgagcg atgttcattc ggtaggggta cagggtgacg ggcgtactta tggttctccc
120
attgtgcttc gcccggtgac gagtgaggac gccatgactg cggactgggc acgtatccca
180
tatgacgtac tggaaaagat ctcgactcgc attacgaatg cgtgtccgca aatcaaccgg
240
gtgggtactcg atatcacatc taaaccgccg gccaccatcg agtgggaatg agccccgtct
300
caccgtgaac atgacatggc cgcaccttt cttggggcgg gccatgccgt gttag
355

```

<210> 88
 <211> 96
 <212> PRT
 <213> Homo sapiens

```

<400> 88
Thr Arg Glu Glu Met Gly Ala Ala Gly Leu Asp Arg Lys Val Trp Gln
1          5          10          15
Cys Pro Val Val Leu Leu Ser Asp Val His Ser Val Gly Val Gln Gly
      20          25          30
Asp Gly Arg Thr Tyr Gly Ser Pro Ile Val Leu Arg Pro Val Thr Ser
      35          40          45
Glu Asp Ala Met Thr Ala Asp Trp Ala Arg Ile Pro Tyr Asp Val Leu
      50          55          60
Glu Lys Ile Ser Thr Arg Ile Thr Asn Ala Cys Pro Gln Ile Asn Arg
65          70          75          80
Val Val Leu Asp Ile Thr Ser Lys Pro Pro Ala Thr Ile Glu Trp Glu
      85          90          95

```

<210> 89
 <211> 351
 <212> DNA
 <213> Homo sapiens

<400> 89

ATTORNEY DOCKET NO.: 15966-543

nacgcgtcaa caccaggcta cgggtgggtat gatcatgata agggctggga cccgcaggag
 60
 agggcaagag ttatataata gacgggtgta tgggtttaag gtatacatgg tgatgagaca
 120
 ggtgattgga taaggccaga ggggtgggag ggggtctgcc cctgctgaag cctgggtggg
 180
 ccaggtctg tgatctggga ccggaacaac acatctgctc tgggcctgct ggatgtggcg
 240
 caagccctgg aacagaacca cagcctcaag tccatgccgc tgccactgaa tgacgtaacc
 300
 caggctcatc gcagccggcc agaactcaca actcgagcgg tccatcagat c
 351

<210> 90

<211> 61

<212> PRT

<213> Homo sapiens

<400> 90

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Ser | Leu | Val | Gly | Pro | Arg | Ser | Val | Ile | Trp | Asp | Arg | Asn | Asn | Thr | Ser |
| 1 | | | | 5 | | | | 10 | | | | | | 15 | |
| Ala | Leu | Gly | Leu | Leu | Asp | Val | Ala | Gln | Ala | Leu | Glu | Gln | Asn | His | Ser |
| | | 20 | | | | | 25 | | | | | | 30 | | |
| Leu | Lys | Ser | Met | Pro | Leu | Pro | Leu | Asn | Asp | Val | Thr | Gln | Ala | His | Arg |
| | 35 | | | | | 40 | | | | | 45 | | | | |
| Ser | Arg | Pro | Glu | Leu | Thr | Thr | Arg | Ala | Val | His | Gln | Ile | | | |
| 50 | | | | | | 55 | | | | | 60 | | | | |

<210> 91

<211> 327

<212> DNA

<213> Homo sapiens

<400> 91

nggtgggtcc tctcgcagga caggtaatcc tgagacctac gcagctccct ggagctctgc
 60
 acgtcgaagc ccagcagggc ctctcgcagg tccctggggc agccagcaca cacaaagtcc
 120
 cggaaggggc tgtagactcc ctgccagcgg ctttcccgga gaaggcacc acgcccagct
 180
 gcctcttgca ggtactgctc ggggtctggtg ggagggcagc cgtgtccagc acaccctgtg
 240
 tgtgcagtcc tctccctgcc ccactgccga acgagccctc cacggtgaag ccattgggga
 300
 acgtgacctt gcccttcccc atgaggt
 327

<210> 92

<211> 107

<212> PRT

<213> Homo sapiens

<400> 92

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Gly | Lys | Gly | Lys | Val | Thr | Phe | Pro | Asn | Gly | Phe | Thr | Val | Glu | Gly |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|

ATTORNEY DOCKET NO.: 15966-543

```

      1           5           10           15
Ser Phe Gly Ser Gly Ala Gly Arg Gly Leu His Thr Gln Gly Val Leu
      20           25           30
Asp Thr Ala Ala Leu Pro Pro Asp Pro Ser Ser Thr Cys Lys Arg Gln
      35           40           45
Leu Gly Val Gly Ala Phe Pro Gly Lys Ala Ala Gly Arg Glu Ser Thr
      50           55           60
Ala Pro Ser Gly Thr Leu Cys Val Leu Ala Ala Pro Gly Thr Cys Arg
      65           70           75           80
Arg Pro Cys Trp Ala Ser Thr Cys Arg Ala Pro Gly Ser Cys Val Gly
      85           90           95
Leu Arg Ile Thr Cys Pro Ala Arg Gly Pro Thr
      100           105

```

<210> 93

<211> 394

<212> DNA

<213> Homo sapiens

<400> 93

```

nccgcgtacg acaagcagta cctcgagggg cgttacgggtg cggacccata cctgagcaac
60
atgctcgaat gggacggcgg acatgagcaa taggccgcca aagcacagcg agaggaaggc
120
ggcgcgctct ggggcctcgc ggatgaagga ggggtggtcg cgcaagagac gccccgcgcc
180
ttttgtaccg ataaatcccg ggcgccacct gatcgtgacc gagggtagca aaacggaacc
240
gctctatttc gaggctatca ggttcggtgt caacaaccgt tatcacggcc agtgggtgac
300
aatggaagtt gtcgttaccg gcaagcatac caggggactt ctgatcgtg cagtcactct
360
ggcgggaagaa agtgccacag gattcactca cgta
394

```

<210> 94

<211> 109

<212> PRT

<213> Homo sapiens

<400> 94

```

Met Leu Ala Gly Asn Asp Asn Phe His Cys His Pro Leu Ala Val Ile
      1           5           10           15
Thr Val Val Asp Thr Gln Pro Asp Ser Leu Glu Ile Glu Arg Phe Arg
      20           25           30
Phe Arg Thr Leu Gly His Asp Gln Val Ala Pro Gly Ile Tyr Arg Tyr
      35           40           45
Lys Arg Arg Gly Ala Ser Leu Ala Arg Pro Pro Leu Leu His Pro Arg
      50           55           60
Gly Arg Arg Ala Arg Arg Leu Pro Leu Ala Val Leu Trp Arg Pro Ile
      65           70           75           80
Ala His Val Arg Arg Pro Ile Arg Ala Cys Cys Ser Gly Met Gly Pro
      85           90           95
His Arg Asn Ala Pro Arg Gly Thr Ala Cys Arg Thr Arg

```

100

105

<210> 95
 <211> 531
 <212> DNA
 <213> Homo sapiens

<400> 95
 ggtacctctt ccaagtacct tctaaatgaa aactcaaga gactgctact caggaaactt
 60
 tgcttggaac ctaaaatgga ctggtcttgg gtgtgtaacc ccggtgaagt tatagcctcc
 120
 ccaaattgag gtgacagaag gaagacaaga ggtgtaagct ggagagggaa gggaagaaat
 180
 cagtggcttt ggccagcctc tgtgccaccc agtacgacag aggagtggga actggccctc
 240
 tggggctctg cttggccata ggcactgcac attgtgccac ctgctcatca cctcctctag
 300
 tctcacactg agcatcggag tacctgttgt gcagacagga aaactgagga gctctgagag
 360
 gctgagcatg gagctcacc ccatgcatag ggtgtgggaa gagggcacag gaggcctcat
 420
 ccattggggga aagggttgag gatggacatg ggtggggaga gggcatagac atcccttctc
 480
 aatctctgtt ccaccacat tcataggag atgagttagg agatgacagc t
 531

<210> 96
 <211> 124
 <212> PRT
 <213> Homo sapiens

<400> 96
 Met Arg Pro Pro Val Pro Ser Ser His Thr Leu Trp His Gly Val Ser
 1 5 10 15
 Ser Met Leu Ser Leu Ser Glu Leu Leu Ser Phe Pro Val Cys Thr Thr
 20 25 30
 Gly Thr Pro Met Leu Ser Val Arg Leu Glu Glu Val Met Ser Arg Trp
 35 40 45
 His Asn Val Gln Cys Leu Trp Pro Ser Arg Ala Pro Glu Gly Gln Phe
 50 55 60
 Pro Leu Leu Cys Arg Thr Gly Trp His Arg Gly Trp Pro Lys Pro Leu
 65 70 75 80
 Ile Ser Ser Leu Pro Ser Pro Ala Tyr Thr Ser Cys Leu Pro Ser Val
 85 90 95
 Thr Ser Ile Trp Gly Gly Tyr Asn Phe Thr Gly Val Thr His Pro Arg
 100 105 110
 Pro Val His Phe Arg Ile Gln Ala Lys Phe Pro Glu
 115 120

<210> 97
 <211> 405
 <212> DNA
 <213> Homo sapiens

<400> 97

ggctcgggcca gtcgaacttc attcccgcctt cgaggggtctt gctgcggatg ggccgtacgc
 60
 tcgcggtgcc ttgcgcgcgg gctggttaggt ggagaagccg cgcgagtacg cgccgtagag
 120
 cgacatcgtg tctgagacgt cgaagctcag gccagcttt ggctccagg cgcgctcggg
 180
 cggctccgcc tcttcggca attgattcag cgcaatcccg gccatcacat gccagcgggt
 240
 gtccaggggtc atgaaatcct gggcataggg gcgcgaggag cgcagcggcg aattggacag
 300
 gcgctcgata ttggcggtga tgtccgaaga cgggaacggg acccgggggg agaagacgtt
 360
 gcccgggaaa agatcccccg acgccatcgt ggtgtcgacc gagat
 405

<210> 98

<211> 122

<212> PRT

<213> Homo sapiens

<400> 98

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Ala | Ser | Gly | Asp | Leu | Phe | Pro | Gly | Asn | Val | Phe | Ser | Pro | Arg | Val |
| 1 | | | | 5 | | | | | 10 | | | | | 15 | |
| Pro | Phe | Pro | Ser | Ser | Asp | Ile | Thr | Pro | Asn | Ile | Glu | Arg | Leu | Ser | Asn |
| | | | 20 | | | | | 25 | | | | | 30 | | |
| Ser | Pro | Leu | Arg | Ser | Ser | Arg | Ala | Tyr | Ala | Gln | Asp | Phe | Met | Thr | Leu |
| | | 35 | | | | 40 | | | | | 45 | | | | |
| Asp | Lys | Arg | Trp | His | Val | Met | Ala | Gly | Ile | Ala | Leu | Asn | Gln | Leu | Pro |
| | 50 | | | 55 | | | | | | 60 | | | | | |
| Gln | Glu | Gly | Gly | Pro | Thr | Glu | Arg | Ala | Trp | Thr | Pro | Lys | Leu | Gly | Leu |
| 65 | | | | 70 | | | | 75 | | | | | | 80 | |
| Ser | Phe | Asp | Val | Ser | Asp | Thr | Met | Ser | Leu | Tyr | Gly | Ala | Tyr | Ser | Arg |
| | | | 85 | | | | | 90 | | | | | 95 | | |
| Gly | Phe | Ser | Thr | Tyr | Gln | Pro | Ala | Arg | Lys | Ala | Pro | Arg | Ala | Tyr | Gly |
| | | | 100 | | | | | 105 | | | | | 110 | | |
| Pro | Ser | Ala | Ala | Arg | Pro | Ser | Lys | Arg | Glu | | | | | | |
| | | 115 | | | | | | 120 | | | | | | | |

<210> 99

<211> 545

<212> DNA

<213> Homo sapiens

<400> 99

acgcgtccgc tcccgatgtc gttgacgagc tgcgctcagc gatgacggta ctcacccatc
 60
 tgccccgacg acccagcaaa cgtccccggc tgttcctcat tgaccacgcc gaccggatcg
 120
 tcgatcccat cactcgggat ttgctggaat ccctgggttcg cgaagccggc gaggctgcgg
 180
 tgatcttggg tgcccagcgt cgcggctcgca tcgattgggt cteccacag atcatccaca
 240

acctggccga acaccatttt gagtcgtcct ctggaggtag tcgatgatga ctgaacgttc
 300
 ccattccacg atcagggttaa ggtggccggc ggtgggtggt ctcgtcctcg ttccgctgct
 360
 ggtgggtcgcc ggattgggtcc gggacgacct ggcataccac cgaccgggtg ggccgggtga
 420
 aagcggccgt cgtcaacgag gacaaggccg tcaagggtgcg tggacaactg gttccgatgg
 480
 gccgccaaact caccgcccgc ttgatggact ctggctcgca caccactgat ggccacaccg
 540
 tcgac
 545

<210> 100
 <211> 101
 <212> PRT
 <213> Homo sapiens

<400> 100
 Met Gly Thr Phe Ser His His Arg Val Pro Pro Glu Asp Asp Ser Lys
 1 5 10 15
 Trp Cys Ser Ala Arg Leu Trp Met Ile Cys Gly Glu Ser Gln Ser Met
 20 25 30
 Arg Pro Arg Arg Trp Ala Pro Lys Ile Thr Ala Ala Ser Pro Ala Ser
 35 40 45
 Arg Thr Arg Asp Ser Ser Lys Ser Arg Val Met Gly Ser Thr Ile Arg
 50 55 60
 Ser Ala Trp Ser Met Arg Asn Ser Arg Gly Arg Leu Leu Gly Arg Arg
 65 70 75 80
 Gly Arg Trp Val Ser Thr Val Ile Ala Glu Arg Ser Ser Ser Thr Thr
 85 90 95
 Ser Gly Ala Asp Ala
 100

<210> 101
 <211> 619
 <212> DNA
 <213> Homo sapiens

<400> 101
 ngcgcgccac agaagtgaag aagtacggtc tcaagcccgt gggagcgttc ttctgtccg
 60
 tcatgcacga tccggacttc gatccgatcc ccatgggtgaa caaggagctt gacgccttcg
 120
 aagctgccgg ggggtgactat ctcacccctcg ccacggattc cggacgcaag ggatacacga
 180
 ccgcccgtcc tcacgaggcc ggcggaaaac gctattacca acctggacca gatccgcgaa
 240
 gtctgcgcca gccgcaacgt caccgcctgt ctacaccccc attggggaac gatgggtccag
 300
 aaccgtgacg aagtgatccg cgtgctcgag aactcctcga tcgggctgtg cctggacact
 360
 ggtcatctgg cctgtggtgg taccgatgtc gttgagctgg tgcgtaagta cgccaaccgc
 420

gtcgacattg tccacgcca agatgtccat aaggagatgg ccgacaagct ttgcctggc
 480
 gagatcacct ggtccgaagg cattcgcgcc gggatgttcg caccatcgg cgacgggtgat
 540
 atcgactttg cagccatcgt gaggctcctt gatgaagccg gggtcgatgg ttattacgtc
 600
 ctagagcagg acatcatga
 619

<210> 102
 <211> 173
 <212> PRT
 <213> Homo sapiens

<400> 102
 Thr Arg Ser Leu Thr Pro Ser Lys Leu Pro Gly Val Thr Ile Ser Ser
 1 5 10 15
 Ser Pro Arg Ile Pro Asp Ala Arg Asp Thr Arg Pro Pro Val Leu Thr
 20 25 30
 Arg Pro Ala Glu Asn Ala Ile Thr Asn Leu Asp Gln Ile Arg Glu Val
 35 40 45
 Cys Ala Ser Arg Asn Val Thr Ala Cys Leu His Pro His Trp Gly Thr
 50 55 60
 Met Val Gln Asn Arg Asp Glu Val Ile Arg Val Leu Glu Asn Ser Ser
 65 70 75 80
 Ile Gly Leu Cys Leu Asp Thr Gly His Leu Ala Cys Gly Gly Thr Asp
 85 90 95
 Val Val Glu Leu Val Arg Lys Tyr Ala Asn Arg Val Asp Ile Val His
 100 105 110
 Ala Lys Asp Val His Lys Glu Met Ala Asp Lys Leu Leu Pro Gly Glu
 115 120 125
 Ile Thr Trp Ser Glu Gly Ile Arg Ala Gly Met Phe Ala Pro Ile Gly
 130 135 140
 Asp Gly Asp Ile Asp Phe Ala Ala Ile Val Arg Leu Leu Asp Glu Ala
 145 150 155 160
 Gly Phe Asp Gly Tyr Tyr Val Leu Glu Gln Asp Ile Met
 165 170

<210> 103
 <211> 321
 <212> DNA
 <213> Homo sapiens

<400> 103
 ncccatgggg gaagacaaca gccatgtggg ggagaccgga gccattgggg ggagaccctt
 60
 gccattgggg ggagaccctt gccgtgggga aagacccttg ccatggggga gaccctggcc
 120
 actgggggga gaccctggcc gctgggggga gaccgagcc attgggggga gaccctggcc
 180
 atggggaaag acccctgcca ttgggggaga ntacctgcca ttgggggaga tccctgccgt
 240
 tggggggaga tccctgctgt tggggggaga ntccctcctg taggggaaga cccctgcagg
 300

agtgggtggg gcgaagaccc c
321

<210> 104

<211> 107

<212> PRT

<213> Homo sapiens

<400> 104

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Xaa | His | Gly | Gly | Arg | Gln | Gln | Pro | Cys | Gly | Gly | Asp | Pro | Ser | His | Trp |
| 1 | | | | 5 | | | | | 10 | | | | | 15 | |
| Gly | Glu | Thr | Pro | Ala | Ile | Gly | Gly | Arg | Pro | Leu | Pro | Trp | Gly | Lys | Thr |
| | | | 20 | | | | | 25 | | | | | 30 | | |
| Pro | Ala | Met | Gly | Gln | Thr | Pro | Ala | Thr | Gly | Gly | Arg | Pro | Leu | Pro | Leu |
| | | 35 | | | | | | 40 | | | | 45 | | | |
| Gly | Gly | Asp | Pro | Ser | His | Trp | Gly | Glu | Thr | Pro | Ala | Met | Gly | Lys | Asp |
| | | 50 | | | | 55 | | | | | 60 | | | | |
| Pro | Cys | His | Trp | Gly | Arg | Xaa | Pro | Ala | Ile | Gly | Gly | Asp | Pro | Cys | Arg |
| 65 | | | | | 70 | | | | | 75 | | | | 80 | |
| Trp | Gly | Glu | Ile | Pro | Ala | Val | Gly | Gly | Arg | Xaa | Pro | Pro | Val | Gly | Glu |
| | | | | 85 | | | | | 90 | | | | | 95 | |
| Asp | Pro | Cys | Arg | Ser | Gly | Trp | Gly | Glu | Asp | Pro | | | | | |
| | | | 100 | | | | | 105 | | | | | | | |

<210> 105

<211> 344

<212> DNA

<213> Homo sapiens

<400> 105

nnntctctcc gaccgcgctcc agatccaccg tggcccgcac gaaccagtcg ttgttgccctt
60
ccgggtcaac gaggggttgg cgcacggctcc actccgtggc gcccgggggtg atgtgcaaca
120
gggcggggccc gcgcgcggcc gggcctgatt ccagcctctc gtgctcgtcc cagtacccat
180
ccagcgcacg gcccagcgg tcggcatccc agccgtggtc gccgtcgagc gcccagggg
240
cctcaatgtc gtcacggcg gccagttcca cccggcggaa catctcgttg cggaccatga
300
cccgaaggc gcgggaattc tcggtcagtt tcggcgggtgc cggc
344

<210> 106

<211> 62

<212> PRT

<213> Homo sapiens

<400> 106

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Cys | Ala | Thr | Gly | Arg | Ala | Arg | Ala | Arg | Pro | Gly | Leu | Ile | Pro | Ala | Ser |
| 1 | | | | 5 | | | | | 10 | | | | | 15 | |
| Arg | Ala | Arg | Pro | Ser | Thr | His | Pro | Ala | His | Arg | Pro | Ser | Gly | Arg | His |
| | | | 20 | | | | | 25 | | | | | 30 | | |
| Pro | Ser | Arg | Gly | Arg | Arg | Arg | Ala | Pro | Pro | Gly | Pro | Gln | Cys | Arg | His |

35 40 45
 Arg Arg Pro Val Pro Pro Gly Gly Thr Ser Arg Cys Gly Pro
 50 55 60

<210> 107
 <211> 549
 <212> DNA
 <213> Homo sapiens

<400> 107
 acgcgttcct cggacgctgc tagtttgagc gcacaggcat tgggcggggc tcacggggga
 60
 gccgcttaat aaccgaccaa catgaaactc aagggtgcc ccttcctagc ggggaccctg
 120
 cacagaccgc aaaataagggtgttttgcctc gccctcctca gttcacgtgg gcaccttgga
 180
 acactgaaga aggcattttc cgaactcact gtcctacgga cttattctcc gcaactgtttt
 240
 cgccctcttc gccctgttct cgtgactgac aggagcagg gtcacaagca ggcagcccg
 300
 gagctctgct cacctggaaa agcatttttg ttagcttaa atgtgaaggc ctcaggcagt
 360
 ggctgttgt cctcctccac atgcgcccac cttcactctt tcatgtgact ggctgtttt
 420
 tgaaggcaag gccctgtca cccttggtta ggccaggat gttctgcacc gaaaatggc
 480
 ctgccctctg cattggatgg ctagctctta ggttggttta ttttagcaaa taagcgttac
 540
 agggtaggc
 549

<210> 108
 <211> 108
 <212> PRT
 <213> Homo sapiens

<400> 108
 Met Lys Leu Lys Gly Cys Pro Phe Leu Ala Gly Thr Leu His Arg Pro
 1 5 10 15
 Glu Asn Lys Gly Phe Cys Ser Ala Leu Leu Ser Ser Arg Gly His Leu
 20 25 30
 Gly Thr Leu Lys Lys Ala Phe Ser Glu Leu Thr Val Leu Arg Thr Tyr
 35 40 45
 Ser Pro His Cys Phe Arg Leu Leu Arg Pro Val Leu Val Thr Asp Arg
 50 55 60
 Ser Arg Gly His Lys Gln Ala Ala Arg Glu Leu Cys Ser Pro Gly Lys
 65 70 75 80
 Ala Phe Leu Cys Ser Leu Asn Val Lys Ala Ser Gly Ser Gly Leu Leu
 85 90 95
 Ser Ser Ser Thr Cys Ala His Leu His Ser Phe Met
 100 105

<210> 109
 <211> 748

<212> DNA

<213> Homo sapiens

<400> 109

nngaattcag atttactttt tgcatttcct tgaatataaa ataggcagta aactaaccta
 60
 aagtcagaag acttatttgt tctggtacct attctgctgt taataaattg ataatgagt
 120
 gatattgggg taagttacca atttacttta cagcccttaa gtaaataatc tgctttcctc
 180
 agcatcatag acttttgaag aggattaatt aagcgcttaa aaaacctgta gactctatta
 240
 cagtcagtga aaggaataat tctctttaca aagtaaatgc agttgtttta ttttagacaa
 300
 gagtgttcta aacttcgtga agagttaagg cttcaacatg aagaggataa gaagtcagca
 360
 atgtctcaac ttttgcagtt gaaagatcga gagaaaaatg cagcaagaga ttcatggcag
 420
 aagaaagtag aagatctctt aaaccagatt tccttgctga aacagaatct ggagatacag
 480
 ctttcccagt ctcagacttc tttgcaacaa ctgcaagccc agtttacgca agaacgacag
 540
 cggcttacgc aagagcttga agaattagag gagcaacatc agcaaagaca caaatcatta
 600
 aaagaagcac atgtccttgc atttcaaact atggaagagg aaaaggaaaa ggagcaaaga
 660
 gctcttgaaa atcatttaca acagaagcat tctgcagagc ttcaatcact aaaagatgca
 720
 cacagagagt caatggaggg cttccgga
 748

<210> 110

<211> 157

<212> PRT

<213> Homo sapiens

<400> 110

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Gln | Leu | Phe | Tyr | Phe | Arg | Gln | Glu | Cys | Ser | Lys | Leu | Arg | Glu | Glu |
| 1 | | | | 5 | | | | | 10 | | | | | 15 | |
| Leu | Arg | Leu | Gln | His | Glu | Glu | Asp | Lys | Lys | Ser | Ala | Met | Ser | Gln | Leu |
| | | 20 | | | | | | 25 | | | | | 30 | | |
| Leu | Gln | Leu | Lys | Asp | Arg | Glu | Lys | Asn | Ala | Ala | Arg | Asp | Ser | Trp | Gln |
| | | 35 | | | | | 40 | | | | | 45 | | | |
| Lys | Lys | Val | Glu | Asp | Leu | Leu | Asn | Gln | Ile | Ser | Leu | Leu | Lys | Gln | Asn |
| | 50 | | | | | 55 | | | | 60 | | | | | |
| Leu | Glu | Ile | Gln | Leu | Ser | Gln | Ser | Gln | Thr | Ser | Leu | Gln | Gln | Leu | Gln |
| 65 | | | | | 70 | | | | | 75 | | | | 80 | |
| Ala | Gln | Phe | Thr | Gln | Glu | Arg | Gln | Arg | Leu | Thr | Gln | Glu | Leu | Glu | Glu |
| | | | 85 | | | | | 90 | | | | | | 95 | |
| Leu | Glu | Glu | Gln | His | Gln | Gln | Arg | His | Lys | Ser | Leu | Lys | Glu | Ala | His |
| | | 100 | | | | | 105 | | | | | | 110 | | |
| Val | Leu | Ala | Phe | Gln | Thr | Met | Glu | Glu | Lys | Glu | Lys | Glu | Gln | Arg | |
| | | 115 | | | | | 120 | | | | | 125 | | | |
| Ala | Leu | Glu | Asn | His | Leu | Gln | Gln | Lys | His | Ser | Ala | Glu | Leu | Gln | Ser |

130 135 140
 Leu Lys Asp Ala His Arg Glu Ser Met Glu Gly Phe Arg
 145 150 155

<210> 111
 <211> 429
 <212> DNA
 <213> Homo sapiens

<400> 111
 gcgcgcccag agagtgcaca atggtgtcag gatatgggag ccaccggaat cattaaccat
 60
 cgtgaaccgc tcgccccgca ggtcaacgac ttcgggatca ccgggttcga cggcattctc
 120
 tcggcttatn nacgccacca gcatncgact ttggctgaga tcatcgcacc gttcggacat
 180
 ctggtcatga tcgacggaac cgactcattc gatctcatgg ccttcaagtc aaagtcgtta
 240
 acggtgacca gcgagtcgat gttcagccgt ccacagttcg cgacgcccga cgtcgccgaa
 300
 caaggccggg cactggccag catcgccgac ctgctcgaga aggggcagat ccgtccgacg
 360
 atgacccgcc atatcgaggg tctgacaacc cagcatgtgc gtgaggccac cgcagccgtc
 420
 gagtccggc
 429

<210> 112
 <211> 143
 <212> PRT
 <213> Homo sapiens

<400> 112
 Ala Arg Pro Glu Ser Ala Gln Trp Cys Gln Asp Met Gly Ala Thr Gly
 1 5 10 15
 Ile Ile Asn His Arg Glu Pro Leu Ala Pro Gln Val Asn Asp Phe Gly
 20 25 30
 Ile Thr Gly Phe Asp Gly Ile Leu Ser Ala Tyr Xaa Arg His Gln His
 35 40 45
 Xaa Thr Leu Ala Glu Ile Ile Ala Pro Phe Gly His Leu Val Met Ile
 50 55 60
 Asp Gly Thr Asp Ser Phe Asp Leu Met Ala Phe Lys Ser Lys Ser Leu
 65 70 75 80
 Thr Val Thr Ser Glu Ser Met Phe Ser Arg Pro Gln Phe Ala Thr Pro
 85 90 95
 Asp Val Ala Glu Gln Gly Arg Ala Leu Ala Ser Ile Ala Asp Leu Val
 100 105 110
 Glu Lys Gly Gln Ile Arg Pro Thr Met Thr Arg His Ile Glu Gly Leu
 115 120 125
 Thr Thr Gln His Val Arg Glu Ala Thr Ala Ala Val Glu Ser Gly
 130 135 140

<210> 113
 <211> 382

<212> DNA

<213> Homo sapiens

<400> 113

gtgcactgga cggatgctgg ggaactgagg cctgcccttc tccactcctg ctgcctctgt
 60
 gtggaccgcc tgccacctgg cttcaacgat gtggacgctc tgtgccgggc gctgtcagct
 120
 gtccacagcc ccaccttctg ccagctggcg tgcggccagg atgggcagct gaagggcttc
 180
 gcggtgctgg agtatgagac ggctgagatg gcggaggagg cacagcagca ggcggacggc
 240
 ctgtccctgg ggggcagcca cctgcgagtc tccttctgcg cccctgggac ccccgggcgc
 300
 agtatgctgg ccgctctcat cggtgcccag gccacggccc tcaatcgggg gcagggagtc
 360
 ctccccgagc ccaacatcct gc
 382

<210> 114

<211> 125

<212> PRT

<213> Homo sapiens

<400> 114

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Leu | Gly | Ser | Gly | Arg | Thr | Pro | Cys | Pro | Arg | Leu | Arg | Ala | Val | Ala |
| 1 | | | | 5 | | | | | 10 | | | | | 15 | |
| Trp | Ala | Thr | Met | Arg | Ala | Ala | Ser | Ile | Leu | Arg | Pro | Gly | Val | Pro | Gly |
| | | | 20 | | | | | 25 | | | | | 30 | | |
| Ala | Gln | Lys | Glu | Thr | Arg | Arg | Trp | Leu | Pro | Pro | Arg | Asp | Arg | Pro | Ser |
| | | 35 | | | | | 40 | | | | | 45 | | | |
| Ala | Cys | Cys | Cys | Ala | Ser | Ser | Ala | Ile | Ser | Ala | Val | Ser | Tyr | Ser | Ser |
| | 50 | | | | | 55 | | | | | 60 | | | | |
| Thr | Ala | Lys | Pro | Phe | Ser | Cys | Pro | Ser | Trp | Pro | His | Ala | Ser | Trp | Gln |
| 65 | | | | | 70 | | | | 75 | | | | | 80 | |
| Lys | Val | Gly | Leu | Trp | Thr | Ala | Asp | Ser | Ala | Arg | His | Arg | Ala | Ser | Thr |
| | | | 85 | | | | | 90 | | | | | | 95 | |
| Ser | Leu | Lys | Pro | Gly | Gly | Arg | Arg | Ser | Thr | Gln | Arg | Gln | Gln | Glu | Trp |
| | | | 100 | | | | | 105 | | | | | | 110 | |
| Arg | Arg | Ala | Gly | Leu | Ser | Ser | Pro | Ala | Ser | Val | Gln | Cys | | | |
| | | | 115 | | | | 120 | | | | | 125 | | | |

<210> 115

<211> 4798

<212> DNA

<213> Homo sapiens

<400> 115

ctcccttcca gtctgtctta agtactgatg ccagagaggt ctgcctaagc tcagtgtgt
 60
 catttcactc cctgtttgga gccatgagtg gccccctgtt gccttcaaga ctgaagccat
 120
 ccccatcctt cctcccacca acctagaggc ttgcttcgt aaatgctggc cctttccttc
 180

atgggctcca ccctctgagt gtgtcatgaa ctcttccact tccttgccct gggtcgtatt
240
gggtccctctg cccgagggtca gagatttggga cgagcccttc tcctccatct tcacagtctc
300
catttaccat tagatgccag gccagacag ttgatgaaaa atgtaagaaa cggacccaag
360
tataagaaaa agctagatct gtccatttta tttctagttt ttggaagaga tgtggataca
420
gagaaatatt actctactat tgaaaaaaat atatatcagt ctgattataa atgtcaactc
480
atactaagtc aatatcaggg aagcagtgtg gagaaggatg tggcaaactg gagattgctg
540
tttccataga aagagggcag ctaatgctta tctccagacc attgttatgc agtaatgtag
600
actccatttc accagattta attcttaagg agaaactgga aacttgtatt attatttgat
660
tattactttt ggcagttaat tttaaaaatc accgttttagg ccaaaacaac atgacatttg
720
tctgggctac aagggactct tagcgactct tagcttgac ttaccttct aatctcaggg
780
ctttttacaa agcccttctc tgaattccct gaagatatgc tagcttgaag gctcactgca
840
tactctcaac tccccagcca ctccctgtt tccctcttat accagttaat actttgatgt
900
tattttttca gtcctattta atcttttaga ctggagggtc cttgagggtg gggcctaatt
960
cttgttttta tcttccccat agcaattatc aagttatctg ctccctttgg cacctaagaa
1020
atattagttg gggtgataac tgaatggctg tccttctgcc ttcgttactg cttttgtgga
1080
tgttattttc tttgacctt attccacctg tgtgcagtgg agaggatggg tgttgatatt
1140
tacagaggaa ggaattaact gagacgcaac aggatgatgg gaacttctca aggtcatgtg
1200
gcaaggaagt caaggaattg gggtttgaac ccaagtcgtc tcagttccat tccattgtcc
1260
tctacccct gccacctctc accttcttcc ctttcacctt tttctgtagc tgaacggaaa
1320
ccaccgctct ttaacatgaa tgcgatgagt gccttatacc acattgcaca gaacgaatcc
1380
cccgtgctcc agtcaggaca ctggtctgag tacttccgga attttgtcga ctctgtctt
1440
cagaaaatcc ctcaagacag accaacctca gaggttctcc tgaagcacg ctttgtgctc
1500
cgggagcggc caccacagt catcatggac ctgatccaga ggaccaagga tgccgtgcgg
1560
gagctggaca acctgcagta ccgcaagatg aagaagatcc tgttccaaga ggcaccaaac
1620
ggccctggtg ccgaggcccc agaggaggaa gaggaggccg agccctacat gcaccgggcc
1680
gggactctga ccagcctcga gagtagccac tcagtgccca gcatgtccat cagcgcctcc
1740
agccagagca gctccgtcaa cagcctagca gatgcctcag acaacgagga agaggaggag
1800

gaggaggagg aagaggagga ggaggaagaa ggccctgaag cccgggagat ggccatgatg
1860
caggaggggg agcacacagt cacctctcac agctccatta tccaccggct gccgggctct
1920
gacaacctat atgatgaccc ctaccagcca gagataaccc ccagccctct ccagccgcct
1980
gcagccccag ctcccacttc caccacctct tccgcccgcc gccgggcta ctgccgtaac
2040
cgagaccact ttgccaccat ccgaaccgcc tccctggtca gccgtcagat ccaggagcat
2100
gagcaggact ctgcgtgctg ggagcagctg agcggctata agcggatgct acgacagcac
2160
cagaagcagc tgctggccct ggagtcacgg ctgaggggtg aacgggagga gcacagtgc
2220
cggtgcagc gggagcttga ggcgagcgg gctggctttg gggcagaggc agaaaagctg
2280
gcccggcggc accaggccat aggtgagaag gaggcacgag ctgccaggc cgaggagcgg
2340
aagtccagc agcacatcct tgggcagcag aagaaggagc tggctgccct gctggaggca
2400
cagaagcggc cctacaaact tcgcaaggaa cagctgaagg aggagctcca ggagaacccc
2460
agcactccca agcgggagaa ggccgagtgg ctgctgcggc agaaggagca gctccagcag
2520
tgccaggcgg aggaggaagc agggctgctg cggcggcagc gccagtactt tgagctgcag
2580
tgtcgccagt acaagcgcaa gatgttgctg gctcggcaca gcctggacca ggacctgctg
2640
cgggaggacc tgaacaagaa gcagaccagc aaggacttgg agtgtgact gctgcttcgg
2700
cagcacgagg ccacgcggga gctggagctg cggcagctcc aggcctgca gcgcacgcgg
2760
gctgagctca cccgcctgca gcaccagacg gagctgggca accagctgga gtacaacaag
2820
cggcgtgagc aagagttgct gcagaagcat gcggcccagg ttcgccagca gcccaagagc
2880
ctcaaagtac gtgcaggcca gcgcccccg ggccttcac tccccattcc tggggctctg
2940
ggccccacca acacaggcac ccctatagaa cagcagccct gctcacctgg ccaggaggca
3000
gtcctggacc aaagaatgct tggcgaggag gaggaagcag ttggagagag aaggattctg
3060
ggaaaggaag gggccacttt ggagcccaag cagcagagga ttctggggga agaatacagga
3120
gcccctagtc ccagtccaca aaaacatggg agcctggttg atgaggaagt ttggggctctg
3180
cctgaggaga tagaggagct taggggtgccc tccctgttac ccaggagag gagcattgct
3240
ggccaggagg aggtgggac gtggagcttg tgggggaagg aggatgagag tcttctggat
3300
gaggagtctg agcttggtg ggtccagggc ccagcactga ctcccgtccc tgaggaggag
3360
gaagaagagg aagagggggc tccgattggg acccctaggg atcctggaga tggttgtcct
3420

tcccccgaca tccctcctga accccctcca acacacctga ggccctgccc tgccagccag
3480
ctccctggac tcctgtccca tggcctcctg gccggcctct cctttgcagt ggggtcctcc
3540
tctggcctcc tgccctcct gctgtgtctg ctgcttccat tgctggcagc ccaggggtggg
3600
ggtggcctgc aggcagcgt gctggccctt gaggtggggc tgggtgggtct gggggcctcc
3660
tacctgctcc tttgtacagc cctgcacctg ccctccagtc ttttctact cctggcccag
3720
ggtaccgcac tggggggcgt cctgggcctg agctggcgcc gaggcctcat ggggtgttccc
3780
ctgggccttg gagctgcctg gctcttagct tggccaggcc tagctctacc tctggtggct
3840
atggcagcgg ggggcagatg ggtgcggcag caaggccccc ggggtgcgccc gggcatatct
3900
cgactctggt tgcgggttct gctgcgcctg tcacccatgg ccttcggggc cctgcagggc
3960
tgtggggctg tgggggaccg ggggctgttt gcaactgtacc ccaaaaccaa caaggatggc
4020
ttccgcagcc gcctgccgt ccctgggccc cggcggcgta atccccgcac cacccaacac
4080
ccattagctc tgttggcaag ggtctgggtc ctgtgcaagg gctggaactg gcgtctggca
4140
cgggcccagc agggtttagc atcccacttg ccccgtggg ccatccacac actggccagc
4200
tggggcctgc ttcgggggtga acgcccacc cgaatcccc ggctactacc acgcagccag
4260
cgccagctgg ggccccctgc ctcccgccag ccaactgccag ggactctagc cgggaggagg
4320
tcacgcaccc gccagtccg ggccctgccc ccctggaggt agctgactcc agcccttcca
4380
gccccaatct agagcattga gcactttatc tcccacgact cagtgaagtt tctccagtcc
4440
ctagtcctct cttttcaccc accttctca gtttgctcac ttaccccagg ccagccctt
4500
cggaccteta gacaggcagc ctctcagct gtggagtcca gcagtcactc tgtgttctcc
4560
tggcgctcct ccctaagtt attgctgttc gcccgctgtg tgtgctcatc ctcacctca
4620
ttgactcagg cctggggcca ggggtgggtg aggggtgggaa gagtcattgt ttttttctcc
4680
tctttgattt tgtttttctg tctcccttcc aacctgtccc ctcccccca ccaaaaaaag
4740
aaaaagacaa acacaaataa aatatctgag cggaactgtg aaaaaaaaaa aaaaaaag
4798

<210> 116

<211> 1062

<212> PRT

<213> Homo sapiens

<400> 116

Met Met Gly Thr Ser Gln Gly His Val Ala Arg Lys Ser Arg Asn Trp

| | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 1 | | | | 5 | | | | | 10 | | | | 15 |
| Gly | Leu | Asn | Pro | Ser | Arg | Leu | Ser | Ser | Ile | Pro | Leu | Ser | Thr |
| | | | 20 | | | | | 25 | | | | 30 | |
| Cys | His | Leu | Ser | Pro | Ser | Ser | Leu | Ser | Pro | Phe | Ser | Val | Ala |
| | | 35 | | | | | 40 | | | | 45 | | |
| Lys | Pro | Pro | Leu | Phe | Asn | Met | Asn | Ala | Met | Ser | Ala | Leu | Tyr |
| | 50 | | | | | 55 | | | | 60 | | | |
| Ala | Gln | Asn | Glu | Ser | Pro | Val | Leu | Gln | Ser | Gly | His | Trp | Ser |
| 65 | | | | | 70 | | | | 75 | | | | 80 |
| Phe | Arg | Asn | Phe | Val | Asp | Ser | Cys | Leu | Gln | Lys | Ile | Pro | Gln |
| | | | 85 | | | | | | 90 | | | | 95 |
| Pro | Thr | Ser | Glu | Val | Leu | Leu | Lys | His | Arg | Phe | Val | Leu | Arg |
| | | | 100 | | | | | 105 | | | | | 110 |
| Pro | Pro | Thr | Val | Ile | Met | Asp | Leu | Ile | Gln | Arg | Thr | Lys | Asp |
| | | 115 | | | | | 120 | | | | | 125 | |
| Arg | Glu | Leu | Asp | Asn | Leu | Gln | Tyr | Arg | Lys | Met | Lys | Lys | Ile |
| | 130 | | | | | 135 | | | | | 140 | | |
| Gln | Glu | Ala | Pro | Asn | Gly | Pro | Gly | Ala | Glu | Ala | Pro | Glu | Glu |
| 145 | | | | | 150 | | | | 155 | | | | 160 |
| Glu | Ala | Glu | Pro | Tyr | Met | His | Arg | Ala | Gly | Thr | Leu | Thr | Ser |
| | | | | 165 | | | | | 170 | | | | 175 |
| Ser | Ser | His | Ser | Val | Pro | Ser | Met | Ser | Ile | Ser | Ala | Ser | Ser |
| | | | 180 | | | | | 185 | | | | 190 | |
| Ser | Ser | Val | Asn | Ser | Leu | Ala | Asp | Ala | Ser | Asp | Asn | Glu | Glu |
| | | 195 | | | | | 200 | | | | 205 | | |
| Glu | Glu | Glu | Glu | Glu | Glu | Glu | Glu | Glu | Glu | Glu | Gly | Pro | Glu |
| | 210 | | | | | 215 | | | | | 220 | | |
| Glu | Met | Ala | Met | Met | Gln | Glu | Gly | Glu | His | Thr | Val | Thr | Ser |
| 225 | | | | | 230 | | | | 235 | | | | 240 |
| Ser | Ile | Ile | His | Arg | Leu | Pro | Gly | Ser | Asp | Asn | Leu | Tyr | Asp |
| | | | 245 | | | | | 250 | | | | | 255 |
| Tyr | Gln | Pro | Glu | Ile | Thr | Pro | Ser | Pro | Leu | Gln | Pro | Pro | Ala |
| | | 260 | | | | | | 265 | | | | 270 | |
| Ala | Pro | Thr | Ser | Thr | Thr | Ser | Ser | Ala | Arg | Arg | Arg | Ala | Tyr |
| | | 275 | | | | | 280 | | | | | 285 | |
| Asn | Arg | Asp | His | Phe | Ala | Thr | Ile | Arg | Thr | Ala | Ser | Leu | Val |
| | 290 | | | | | 295 | | | | | 300 | | |
| Gln | Ile | Gln | Glu | His | Glu | Gln | Asp | Ser | Ala | Leu | Arg | Glu | Gln |
| 305 | | | | | 310 | | | | 315 | | | | 320 |
| Gly | Tyr | Lys | Arg | Met | Arg | Arg | Gln | His | Gln | Lys | Gln | Leu | Leu |
| | | | 325 | | | | | | 330 | | | | 335 |
| Glu | Ser | Arg | Leu | Arg | Gly | Glu | Arg | Glu | Glu | His | Ser | Ala | Arg |
| | | 340 | | | | | 345 | | | | | 350 | |
| Arg | Glu | Leu | Glu | Ala | Gln | Arg | Ala | Gly | Phe | Gly | Ala | Glu | Ala |
| | | 355 | | | | | 360 | | | | | 365 | |
| Leu | Ala | Arg | Arg | His | Gln | Ala | Ile | Gly | Glu | Lys | Glu | Ala | Arg |
| | 370 | | | | | 375 | | | | | 380 | | |
| Gln | Ala | Glu | Glu | Arg | Lys | Phe | Gln | Gln | His | Ile | Leu | Gly | Gln |
| 385 | | | | | 390 | | | | 395 | | | | 400 |
| Lys | Glu | Leu | Ala | Ala | Leu | Leu | Glu | Ala | Gln | Lys | Arg | Thr | Tyr |
| | | | 405 | | | | | | 410 | | | | 415 |
| Arg | Lys | Glu | Gln | Lys | Glu | Glu | Leu | Gln | Glu | Asn | Pro | Ser | Thr |
| | | 420 | | | | | 425 | | | | 430 | | |
| Lys | Arg | Glu | Lys | Ala | Glu | Trp | Leu | Leu | Arg | Gln | Lys | Glu | Gln |

435 440 445
 Gln Cys Gln Ala Glu Glu Glu Ala Gly Leu Leu Arg Arg Gln Arg Gln
 450 455 460
 Tyr Phe Glu Leu Gln Cys Arg Gln Tyr Lys Arg Lys Met Leu Leu Ala
 465 470 475 480
 Arg His Ser Leu Asp Gln Asp Leu Leu Arg Glu Asp Leu Asn Lys Lys
 485 490 495
 Gln Thr Gln Lys Asp Leu Glu Cys Ala Leu Leu Leu Arg Gln His Glu
 500 505 510
 Ala Thr Arg Glu Leu Glu Leu Arg Gln Leu Gln Ala Val Gln Arg Thr
 515 520 525
 Arg Ala Glu Leu Thr Arg Leu Gln His Gln Thr Glu Leu Gly Asn Gln
 530 535 540
 Leu Glu Tyr Asn Lys Arg Arg Glu Gln Glu Leu Arg Gln Lys His Ala
 545 550 555 560
 Ala Gln Val Arg Gln Gln Pro Lys Ser Leu Lys Val Arg Ala Gly Gln
 565 570 575
 Arg Pro Pro Gly Leu Pro Leu Pro Ile Pro Gly Ala Leu Gly Pro Pro
 580 585 590
 Asn Thr Gly Thr Pro Ile Glu Gln Gln Pro Cys Ser Pro Gly Gln Glu
 595 600 605
 Ala Val Leu Asp Gln Arg Met Leu Gly Glu Glu Glu Glu Ala Val Gly
 610 615 620
 Glu Arg Arg Ile Leu Gly Lys Glu Gly Ala Thr Leu Glu Pro Lys Gln
 625 630 635 640
 Gln Arg Ile Leu Gly Glu Glu Ser Gly Ala Pro Ser Pro Ser Pro Gln
 645 650 655
 Lys His Gly Ser Leu Val Asp Glu Glu Val Trp Gly Leu Pro Glu Glu
 660 665 670
 Ile Glu Glu Leu Arg Val Pro Ser Leu Val Pro Gln Glu Arg Ser Ile
 675 680 685
 Val Gly Gln Glu Glu Ala Gly Thr Trp Ser Leu Trp Gly Lys Glu Asp
 690 695 700
 Glu Ser Leu Leu Asp Glu Glu Phe Glu Leu Gly Trp Val Gln Gly Pro
 705 710 715 720
 Ala Leu Thr Pro Val Pro Glu Glu Glu Glu Glu Glu Glu Gly Ala
 725 730 735
 Pro Ile Gly Thr Pro Arg Asp Pro Gly Asp Gly Cys Pro Ser Pro Asp
 740 745 750
 Ile Pro Pro Glu Pro Pro Pro Thr His Leu Arg Pro Cys Pro Ala Ser
 755 760 765
 Gln Leu Pro Gly Leu Leu Ser His Gly Leu Leu Ala Gly Leu Ser Phe
 770 775 780
 Ala Val Gly Ser Ser Ser Gly Leu Leu Pro Leu Leu Leu Leu Leu Leu
 785 790 795 800
 Leu Pro Leu Leu Ala Ala Gln Gly Gly Gly Gly Leu Gln Ala Ala Leu
 805 810 815
 Leu Ala Leu Glu Val Gly Leu Val Gly Leu Gly Ala Ser Tyr Leu Leu
 820 825 830
 Leu Cys Thr Ala Leu His Leu Pro Ser Ser Leu Phe Leu Leu Leu Ala
 835 840 845
 Gln Gly Thr Ala Leu Gly Ala Val Leu Gly Leu Ser Trp Arg Arg Gly
 850 855 860
 Leu Met Gly Val Pro Leu Gly Leu Gly Ala Ala Trp Leu Leu Ala Trp

| 865 | | | | | | | | | | 870 | | | | | | | | | | 875 | | | | | | | | | | 880 | | | | | | | | | |
|------|-----|-----|-----|-----|-----|-----|-----|-----|-----|------|-----|-----|-----|-----|-----|--|--|--|--|------|--|--|--|--|--|--|--|--|--|-----|--|--|--|--|--|--|--|--|--|
| Pro | Gly | Leu | Ala | Leu | Pro | Leu | Val | Ala | Met | Ala | Ala | Gly | Gly | Arg | Trp | | | | | | | | | | | | | | | | | | | | | | | | |
| 885 | | | | | | | | | | 890 | | | | | | | | | | 895 | | | | | | | | | | | | | | | | | | | |
| Val | Arg | Gln | Gln | Gly | Pro | Arg | Val | Arg | Arg | Gly | Ile | Ser | Arg | Leu | Trp | | | | | | | | | | | | | | | | | | | | | | | | |
| 900 | | | | | | | | | | 905 | | | | | | | | | | 910 | | | | | | | | | | | | | | | | | | | |
| Leu | Arg | Val | Leu | Leu | Arg | Leu | Ser | Pro | Met | Ala | Phe | Arg | Ala | Leu | Gln | | | | | | | | | | | | | | | | | | | | | | | | |
| 915 | | | | | | | | | | 920 | | | | | | | | | | 925 | | | | | | | | | | | | | | | | | | | |
| Gly | Cys | Gly | Ala | Val | Gly | Asp | Arg | Gly | Leu | Phe | Ala | Leu | Tyr | Pro | Lys | | | | | | | | | | | | | | | | | | | | | | | | |
| 930 | | | | | | | | | | 935 | | | | | | | | | | 940 | | | | | | | | | | | | | | | | | | | |
| Thr | Asn | Lys | Asp | Gly | Phe | Arg | Ser | Arg | Leu | Pro | Val | Pro | Gly | Pro | Arg | | | | | | | | | | | | | | | | | | | | | | | | |
| 945 | | | | | | | | | | 950 | | | | | | | | | | 955 | | | | | | | | | | | | | | | | | | | |
| Arg | Arg | Asn | Pro | Arg | Thr | Thr | Gln | His | Pro | Leu | Ala | Leu | Leu | Ala | Arg | | | | | | | | | | | | | | | | | | | | | | | | |
| 965 | | | | | | | | | | 970 | | | | | | | | | | 975 | | | | | | | | | | | | | | | | | | | |
| Val | Trp | Val | Leu | Cys | Lys | Gly | Trp | Asn | Trp | Arg | Leu | Ala | Arg | Ala | Ser | | | | | | | | | | | | | | | | | | | | | | | | |
| 980 | | | | | | | | | | 985 | | | | | | | | | | 990 | | | | | | | | | | | | | | | | | | | |
| Gln | Gly | Leu | Ala | Ser | His | Leu | Pro | Pro | Trp | Ala | Ile | His | Thr | Leu | Ala | | | | | | | | | | | | | | | | | | | | | | | | |
| 995 | | | | | | | | | | 1000 | | | | | | | | | | 1005 | | | | | | | | | | | | | | | | | | | |
| Ser | Trp | Gly | Leu | Leu | Arg | Gly | Glu | Arg | Pro | Thr | Arg | Ile | Pro | Arg | Leu | | | | | | | | | | | | | | | | | | | | | | | | |
| 1010 | | | | | | | | | | 1015 | | | | | | | | | | 1020 | | | | | | | | | | | | | | | | | | | |
| Leu | Pro | Arg | Ser | Gln | Arg | Gln | Leu | Gly | Pro | Pro | Ala | Ser | Arg | Gln | Pro | | | | | | | | | | | | | | | | | | | | | | | | |
| 1025 | | | | | | | | | | 1030 | | | | | | | | | | 1035 | | | | | | | | | | | | | | | | | | | |
| Leu | Pro | Gly | Thr | Leu | Ala | Gly | Arg | Arg | Ser | Arg | Thr | Arg | Gln | Ser | Arg | | | | | | | | | | | | | | | | | | | | | | | | |
| 1045 | | | | | | | | | | 1050 | | | | | | | | | | 1055 | | | | | | | | | | | | | | | | | | | |
| Ala | Leu | Pro | Pro | Trp | Arg | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 1060 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```
<210> 117
<211> 471
<212> DNA
<213> Homo sapiens
```

```
<400> 117
naccgcttga cgatctgtct ggctgggtga gtgatctgcg ctgtgggtgt cgtcgatgac
60
ctgctcgacc ttctctgctt ggccaaggca gctggccagg tattagcggc cggcatcgtc
120
gtcacgggcg gagtgcgaat gttttggatc ccgctgccga actccatcat tgctttgggg
180
acgcctactt cgatcttggg gacggtgttc ttcatgtgtg tgtgcgcaa tgcggtgaat
240
ttcattgatg gacttgacgg cctggcatcc ggtgtgggtg ccatcgggtc cttggctttc
300
ttctcataca cctacctgct ggctcacgaa caggactttg ttgttgcgac gactaccagt
360
ctcattacgg ctgcgacggc gggcgccgtg ctcggttttt tgccccacaa ctggcatccg
420
gcgaggatgt tcatgggtga ttccggagct ctgctacttg gcttattgct a
471
```

```
<210> 118
<211> 157
<212> PRT
<213> Homo sapiens
```


<400> 118

```

Xaa Ala Leu Thr Ile Cys Leu Ala Gly Val Val Ile Cys Ala Val Gly
 1           5           10           15
Val Val Asp Asp Leu Leu Asp Leu Pro Ala Leu Ala Lys Ala Ala Gly
      20           25           30
Gln Val Leu Ala Ala Gly Ile Val Val Thr Gly Gly Val Arg Met Phe
      35           40           45
Trp Ile Pro Leu Pro Asn Ser Ile Ile Ala Leu Gly Thr Pro Thr Ser
      50           55           60
Ile Leu Val Thr Val Phe Phe Ile Val Leu Cys Ala Asn Ala Val Asn
      65           70           75           80
Phe Ile Asp Gly Leu Asp Gly Leu Ala Ser Gly Val Val Ala Ile Gly
      85           90           95
Ser Leu Ala Phe Phe Ser Tyr Thr Tyr Leu Leu Ala His Glu Gln Asp
      100          105          110
Phe Val Val Ala Thr Thr Thr Ser Leu Ile Thr Ala Ala Thr Ala Gly
      115          120          125
Ala Cys Leu Gly Phe Leu Pro His Asn Trp His Pro Ala Arg Met Phe
      130          135          140
Met Gly Asp Ser Gly Ala Leu Leu Leu Gly Leu Leu Leu
      145          150          155

```

<210> 119

<211> 302

<212> DNA

<213> Homo sapiens

<400> 119

```

ntcaaacatg agcagtcgtg gcggccgagg ccgcggtggc tattatcgcg agctttatgg
60
tagccgaggt cgaggcagta aatctaataa aacttttcgca aaaaattcgg atgtctactc
120
tcagaaaaag actcgaacag tacgaggcac ctccgaagat ttagcacgat cgctccataa
180
gcttcatatg cgcccgtaac ctgcgtatca tgacattgag ggtatgtggg ctttcccagc
240
ctttactttt tatctggatc atgcacaagc agaccatac gctgccccaa ataaggcacg
300
cn
302

```

<210> 120

<211> 98

<212> PRT

<213> Homo sapiens

<400> 120

```

Met Ser Ser Arg Gly Gly Arg Gly Gly Tyr Tyr Arg Glu Leu
 1           5           10           15
Tyr Gly Ser Arg Gly Arg Gly Ser Lys Ser Asn Glu Thr Phe Ala Lys
      20           25           30
Asn Ser Asp Val Tyr Ser Gln Lys Lys Thr Arg Thr Val Arg Gly Thr
      35           40           45
Ser Glu Asp Leu Ala Arg Ser Leu His Lys Leu His Met Arg Pro Tyr

```

50 55 60
 Pro Ala Tyr His Asp Ile Glu Gly Met Trp Ala Phe Pro Ala Phe Thr
 65 70 75 80
 Phe Tyr Leu Asp His Ala Gln Ala Asp Pro Tyr Ala Ala Pro Asn Lys
 85 90 95
 Ala Arg

<210> 121
 <211> 318
 <212> DNA
 <213> Homo sapiens

<400> 121
 ngcatggggg gccctgggac cgcacttggtg cccctttttt ttttagggaa aaaattgagc
 60
 cctaaaggat ttgccgcatt acaggaaagt tttttggtaa gtttgggggtt gtttctgtgc
 120
 tgtgtgagaa ggagtagaag cagctccagt agagtggggc ttttcatttt tatccagagg
 180
 aaattttagt gctgtggcta ttacttcctt tttttctttt tttttttttg ttttagagaca
 240
 gagtctgnct ctgtcgccag gctggagtga agtggcacga tctcagctca ctgcaacctc
 300
 tgcctcccag gttcaagc
 318

<210> 122
 <211> 89
 <212> PRT
 <213> Homo sapiens

<400> 122
 Xaa Met Gly Gly Pro Gly Thr Ala Leu Val Pro Leu Phe Phe Leu Gly
 1 5 10 15
 Lys Lys Leu Ser Pro Lys Gly Phe Ala Ala Leu Gln Glu Ser Phe Leu
 20 25 30
 Val Ser Leu Gly Leu Phe Leu Cys Cys Val Arg Arg Ser Arg Ser Ser
 35 40 45
 Ser Ser Arg Val Gly Leu Phe Ile Phe Ile Gln Arg Lys Phe Val Gly
 50 55 60
 Cys Gly Tyr Tyr Phe Leu Phe Phe Leu Phe Phe Phe Cys Leu Glu Thr
 65 70 75 80
 Glu Ser Xaa Ser Val Ala Arg Leu Glu
 85

<210> 123
 <211> 338
 <212> DNA
 <213> Homo sapiens

<400> 123
 acgcgtctag ggtagaaatc aactccagta actgtcatc aacctcagca atgctgggga
 60

cgggcagagg cagggcagct gtgtgccaca ttctgccag ggctggtcag gccccggctc
 120
 tcaccactcc tcctccctgc tttgaacctg tggaacaaag ggcccctgca ccccaactca
 180
 ttctcttttg ccacataagg gcctcaagtc atgctgtccc ctctgcctgg gttgcttttt
 240
 ctccctctgc ttgggtcact gttcacacca ctggccactt tcctcagga agggccctca
 300
 ctgcccacac acctaaacat gccccctgct cctccata
 338

<210> 124

<211> 96

<212> PRT

<213> Homo sapiens

<400> 124

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Leu | Gly | Thr | Gly | Arg | Gly | Arg | Ala | Ala | Val | Cys | His | Ile | Pro | Ala |
| 1 | | | | 5 | | | | 10 | | | | | 15 | | |
| Arg | Ala | Gly | Gln | Ala | Pro | Ala | Leu | Thr | Thr | Pro | Pro | Pro | Cys | Phe | Glu |
| | | 20 | | | | | 25 | | | | | 30 | | | |
| Pro | Val | Glu | Gln | Arg | Ala | Pro | Ala | Pro | Gln | Leu | Ile | Pro | Leu | Cys | His |
| | 35 | | | | | 40 | | | | 45 | | | | | |
| Ile | Arg | Ala | Ser | Ser | His | Ala | Val | Pro | Ser | Ala | Trp | Val | Ala | Phe | Ser |
| 50 | | | | | 55 | | | | 60 | | | | | | |
| Pro | Ser | Ala | Trp | Val | Thr | Val | His | Thr | Thr | Gly | His | Phe | Pro | Gln | Gly |
| 65 | | | | 70 | | | | 75 | | | | | 80 | | |
| Arg | Ala | Leu | Thr | Ala | His | Thr | Pro | Lys | His | Ala | Pro | Cys | Ser | Ser | Ile |
| | | | 85 | | | | | 90 | | | | | 95 | | |

<210> 125

<211> 280

<212> DNA

<213> Homo sapiens

<400> 125

ccatggacct ggccagccac catcacctgc ctctgcctc acccaccctg ggtgcctgcc
 60
 ggcaaggatt ggagggcaga ctgctggagc gtgagaccag gccaatctgt ctttctggga
 120
 accttcagcc tccaactgga gctgactgtc aactttcggg tgagaagtca cttttctgca
 180
 ttcccaccac actatctatc tgtgcaatac ggcagcgtga cagcactcac cttattgagg
 240
 gcttctgctg tcctggccca ttctggatag gcctgatcta
 280

<210> 126

<211> 92

<212> PRT

<213> Homo sapiens

<400> 126

Met Asp Leu Ala Ser His His His Leu Pro Pro Ala Ser Pro Thr Leu

```

      1             5             10             15
Gly Ala Cys Arg Gln Gly Leu Glu Gly Arg Leu Leu Glu Arg Glu Thr
      20             25             30
Arg Pro Ile Cys Leu Ser Gly Asn Leu Gln Pro Pro Thr Gly Ala Asp
      35             40             45
Cys Gln Leu Ser Gly Glu Lys Ser Leu Phe Cys Ile Pro Thr Thr Leu
      50             55             60
Ser Ile Cys Ala Ile Arg Gln Arg Asp Ser Thr His Leu Ile Glu Gly
      65             70             75             80
Phe Cys Cys Pro Gly Pro Phe Trp Ile Gly Leu Ile
      85             90

```

<210> 127
 <211> 444
 <212> DNA
 <213> Homo sapiens

```

<400> 127
cgcgatgatcg ccgtggcgga gggccgcggc gccgactcga tcgcccagct gacaaccgag
60
ctgcaaagcc gtcactgccc tgcggagcag atcacgtccg tcagcatcga catgtcgcca
120
gcgttcatca ggggctgcgc cgagcacctg cccaacgcgc gcgtcacctt cgacaagttc
180
cacgtcatcg ggcacgcaa tgcggccgtg gacaggatgc gccgcatcga gcagcgcagc
240
gacaagtccc tcaaggggat gcgctggtcg ctgctgaaga accgcgccag cctcaagccc
300
gaggctgccg ccgatctgga tgccctgac gccaggatgg ccactgtgcg caccgcgcgc
360
gcctgggtct acaaggagca gctgcgcgag atcctcgcgc gcaagcagat caacgtggca
420
cgcgacatgc tcaagcactg gtgc
444

```

<210> 128
 <211> 148
 <212> PRT
 <213> Homo sapiens

```

<400> 128
Arg Val Ile Ala Val Ala Glu Gly Arg Gly Ala Asp Ser Ile Ala Gln
1             5             10             15
Leu Thr Thr Glu Leu Gln Ser Arg His Cys Pro Ala Glu Gln Ile Thr
      20             25             30
Ser Val Ser Ile Asp Met Ser Pro Ala Phe Ile Arg Gly Cys Ala Glu
      35             40             45
His Leu Pro Asn Ala Arg Val Thr Phe Asp Lys Phe His Val Ile Gly
      50             55             60
His Ala Asn Ala Ala Val Asp Arg Met Arg Arg Ile Glu Gln Arg Ser
      65             70             75             80
Asp Lys Ser Leu Lys Gly Met Arg Trp Ser Leu Leu Lys Asn Arg Ala
      85             90             95
Ser Leu Lys Pro Glu Ala Ala Ala Asp Leu Asp Ala Leu Ile Ala Arg

```

100 105 110
 Met Ala Thr Val Arg Thr Ala Arg Ala Trp Val Tyr Lys Glu Gln Leu
 115 120 125
 Arg Glu Ile Leu Ala Arg Lys Gln Ile Asn Val Ala Arg Asp Met Leu
 130 135 140
 Lys His Trp Cys
 145

<210> 129
 <211> 291
 <212> DNA
 <213> Homo sapiens

<400> 129
 gaggaggac gtaccgtccc cggtatagcc aagctcgaga agccgcaagc tatcgagaac
 60
 ttggacgaga ttattgacgt ctttgacgcc gtcattggtg cccgtggcga tatggccgtc
 120
 gagtgccgc tgcaggaagt tccgctgac caaaagcaga tcatcgagaa ggctcgttta
 180
 caggctaagc ccgtcattgt ggccaccag atgcttgagt cgatgatcca cgctccccgt
 240
 ccgaccgcg ctgaggccgc cgacgtcgcg aacgccatcc ttgacggcgc g
 291

<210> 130
 <211> 97
 <212> PRT
 <213> Homo sapiens

<400> 130
 Glu Glu Gly Arg Thr Val Pro Val Ile Ala Lys Leu Glu Lys Pro Gln
 1 5 10 15
 Ala Ile Glu Asn Leu Asp Glu Ile Ile Asp Val Phe Asp Ala Val Met
 20 25 30
 Val Ala Arg Gly Asp Met Ala Val Glu Cys Pro Leu Glu Glu Val Pro
 35 40 45
 Leu Ile Gln Lys Gln Ile Ile Glu Lys Ala Arg Leu Gln Ala Lys Pro
 50 55 60
 Val Ile Val Ala Thr Gln Met Leu Glu Ser Met Ile His Ala Pro Arg
 65 70 75 80
 Pro Thr Arg Ala Glu Ala Ala Asp Val Ala Asn Ala Ile Leu Asp Gly
 85 90 95
 Ala

<210> 131
 <211> 416
 <212> DNA
 <213> Homo sapiens

<400> 131
 tccggagcgt ccgtggccct catgggtgtg tcagcgtggt tgctgtctcg ggccgcagag
 60

attccaccgg tgctctacct ggaggccgca gccgtcgggg ttcgattctt cggcatctcc
 120
 cgcggtgtct tccgctacgc cgaacgtctg gtaggccacg acctggctct gcggatgcag
 180
 ggggcattgc ggatgcgggt ctacgaccgg ctgtcacgta ccnaccctgc tgggnnacgt
 240
 cgccgggggtg acctgctggg acgggttact gccgacgtcg acgcggtgtt ggacatgggc
 300
 gtgcgggtga tcgttccggc gtgcgcgtca agcctcgtca tcattggcac cacggtcctt
 360
 ctttgtccga gagaaggttg agttttctta gccggattcc aacacagcct gggggg
 416

<210> 132

<211> 126

<212> PRT

<213> Homo sapiens

<400> 132

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Ser | Gly | Ala | Ser | Val | Ala | Leu | Met | Gly | Val | Ser | Ala | Trp | Leu | Leu | Ser |
| 1 | | | | 5 | | | | | 10 | | | | | 15 | |
| Arg | Ala | Ala | Glu | Ile | Pro | Pro | Val | Leu | Tyr | Leu | Glu | Ala | Ala | Ala | Val |
| | | | 20 | | | | | 25 | | | | | 30 | | |
| Gly | Val | Arg | Phe | Phe | Gly | Ile | Ser | Arg | Gly | Val | Phe | Arg | Tyr | Ala | Glu |
| | | 35 | | | | | 40 | | | | 45 | | | | |
| Arg | Leu | Val | Gly | His | Asp | Leu | Ala | Leu | Arg | Met | Gln | Gly | Ala | Leu | Arg |
| | 50 | | | | | 55 | | | | 60 | | | | | |
| Met | Arg | Val | Tyr | Asp | Arg | Leu | Ser | Arg | Thr | Xaa | Pro | Ala | Gly | Xaa | Arg |
| 65 | | | | 70 | | | | | 75 | | | | | 80 | |
| Arg | Arg | Gly | Asp | Leu | Leu | Val | Arg | Val | Thr | Ala | Asp | Val | Asp | Ala | Val |
| | | | 85 | | | | | 90 | | | | | 95 | | |
| Leu | Asp | Met | Val | Val | Arg | Val | Ile | Val | Pro | Ala | Cys | Ala | Ser | Ser | Leu |
| | | 100 | | | | | | 105 | | | | | 110 | | |
| Val | Ile | Ile | Gly | Thr | Thr | Val | Leu | Leu | Cys | Pro | Arg | Glu | Gly | | |
| | | 115 | | | | | 120 | | | | | | 125 | | |

<210> 133

<211> 327

<212> DNA

<213> Homo sapiens

<400> 133

gccgttgcta tcgctgctgg tatgcgtgca gacgtcactg tttttgatat caatatcgct
 60
 gcggtgaaga gactcgccga catctaccag ggctgtgttc acacagtagt atccaccgc
 120
 gccgaaattg cgaaggcgct agaaaccgct gacgttggtga tcggttctgt ccttattccg
 180
 ggtagttcta ccccgaagct tggtactacc gatatgggtg ctcacatgca gcctgggtct
 240
 gttcttattg atattgctat agaccaaggc ggctgcttcg aggattcgca cccaccact
 300
 tacgatgacc ccactttcac tgtgcac
 327

<210> 134
 <211> 109
 <212> PRT
 <213> Homo sapiens

<400> 134
 Ala Val Ala Ile Ala Ala Gly Met Arg Ala Asp Val Thr Val Phe Asp
 1 5 10 15
 Ile Asn Ile Ala Ala Leu Lys Arg Leu Ala Asp Ile Tyr Gln Gly Arg
 20 25 30
 Val His Thr Val Val Ser Thr Arg Ala Glu Ile Ala Lys Ala Leu Glu
 35 40 45
 Thr Ala Asp Val Val Ile Gly Ser Val Leu Ile Pro Gly Ser Ser Thr
 50 55 60
 Pro Lys Leu Val Thr Thr Asp Met Val Ala His Met Gln Pro Gly Ser
 65 70 75 80
 Val Leu Ile Asp Ile Ala Ile Asp Gln Gly Gly Cys Phe Glu Asp Ser
 85 90 95
 His Pro Thr Thr Tyr Asp Asp Pro Thr Phe Thr Val His
 100 105

<210> 135
 <211> 560
 <212> DNA
 <213> Homo sapiens

<400> 135
 taagatgtgg tcctgccttg ttctgaagg ggctgcagct ctgatggaaa atacagggat
 60
 ttacactcag ggctacagcc acggggggct gaggcccaag gctgcaatct cgggggaagg
 120
 ggaagttggc ttttcttggt ggattggaaa catcctcttg gaggcaaaga ctttctcttg
 180
 atcttacaga cttcccgga ttttagatt agaattattgg gggcaaagga ggctgtcttg
 240
 ttttaaagca atgtacata gacacagtgg ggaagacctg gttcgacggc agataagcag
 300
 tgggtgatgg gcttgaggag gagagtcagg gcaaagtcta agactgagca gaaaggaatt
 360
 ccccatctc ccatggataa gtacgttcta gaacattctc ttgggtcta atactctgaa
 420
 atgacatctt gtcttcatgc tcgagagaga attacttcac tggctccact tggagtgcca
 480
 gtgttcagac accaagcctg actgggaggg ttccgttttc ttaacacctt cccaccgccg
 540
 attccaagt cccacgcgt
 560

<210> 136
 <211> 100
 <212> PRT
 <213> Homo sapiens

<400> 136

```

Met Trp Ser Cys Pro Val Pro Glu Gly Ala Ala Ala Leu Met Glu Asn
 1             5             10             15
Thr Gly Ile Tyr Thr Gln Gly Tyr Ser His Gly Gly Leu Arg Pro Lys
      20             25             30
Ala Ala Ile Ser Gly Glu Gly Glu Val Gly Phe Ser Trp Trp Ile Gly
      35             40             45
Asn Ile Leu Leu Glu Ala Lys Thr Phe Pro Gly Ser Tyr Arg Leu Pro
      50             55             60
Gly Ile Phe Arg Leu Glu Tyr Trp Gly Gln Arg Arg Leu Ser Cys Phe
      65             70             75             80
Lys Ala Met Leu His Arg His Ser Gly Glu Asp Leu Val Arg Arg Gln
      85             90             95
Ile Ser Ser Gly
      100

```

<210> 137

<211> 429

<212> DNA

<213> Homo sapiens

<400> 137

```

accggttgga tggcctgcag gccaaagcgt tcctgcaaac tcagcaggcc ttcagcgcaa
60
gaggcaaaca gctggtcgcg cacctgcttg aggtccaccg attgcgcata gcccttgagc
120
aaggcgcgcc agttggtttt gtcggccact tggctgcgga acaggtcttc gacaaaaccg
180
gactgctggc gggtcgcaac gcgcatgata ggcagcgccg ggctggcgcc ctggctcgagc
240
cagcgctcgc gcagttgggt ggcccgggtg ataccgacct tgatccccga cgaattggcc
300
aggtacacca catggtcggc catgcagaat gtttcgcccc agccgggata acggcaagtg
360
ccggcgctgt aatggcaacg ttcggggctc atgatgcaca ggtcacactg ggccagcttg
420
gtcatgccc
429

```

<210> 138

<211> 141

<212> PRT

<213> Homo sapiens

<400> 138

```

Met Thr Lys Leu Ala Gln Cys Asp Leu Cys Ile Met Ser Pro Glu Arg
 1             5             10             15
Cys His Tyr Asp Ala Gly Thr Cys Arg Asp Pro Gly Trp Gly Glu Thr
      20             25             30
Phe Cys Met Thr Asp His Val Val Tyr Leu Ala Asn Ser Ser Gly Ile
      35             40             45
Lys Val Gly Ile Thr Arg Ala Thr Gln Leu Pro Thr Arg Trp Leu Asp
      50             55             60
Gln Gly Ala Ser Gln Ala Leu Pro Ile Met Arg Val Ala Thr Arg Gln

```



```

65              70              75              80
Gln Ser Gly Phe Val Glu Asp Leu Phe Arg Ser Gln Val Ala Asp Lys
              85              90              95
Thr Asn Trp Arg Ala Leu Leu Lys Gly Asp Ala Gln Ser Val Asp Leu
              100              105              110
Lys Gln Val Arg Asp Gln Leu Phe Ala Ser Cys Ala Glu Gly Leu Leu
              115              120              125
Ser Leu Gln Glu Arg Phe Gly Leu Gln Ala Ile Gln Pro
              130              135              140

```

<210> 139
 <211> 341
 <212> DNA
 <213> Homo sapiens

```

<400> 139
acgcgtcgtt tgaaggcttg atccgcacgt ccaattcgct ttgcgccaat ggcgcgcagc
60
ttgtgaacag cagaatcaag ccgctggtaa atcttctctg gagcttcata ggcggggatg
120
ctacacgagc tggggagaca ctttgaaccc ggaattgtct gaataattct gtctcaaacc
180
tttgcagcct gtaacgactg agggttcgga tggaaaaaca catgctccag gatggggaccg
240
acggccactt caccgatctc ttcatagccc tggcgtttgt agaaatccag gtagcgcgaa
300
tcgccagcgt cgagcacgac gctgatgag tgcgggtcat t
341

```

<210> 140
 <211> 113
 <212> PRT
 <213> Homo sapiens

```

<400> 140
Met Thr Arg Thr His Gln Ala Ser Cys Ser Thr Leu Ala Ile Arg Ala
1              5              10              15
Thr Trp Ile Ser Thr Asn Ala Arg Ala Met Lys Arg Ser Val Lys Trp
              20              25              30
Pro Ser Val Pro Ser Trp Ser Met Cys Phe Ser Ile Arg Thr Leu Ser
              35              40              45
Arg Tyr Arg Leu Gln Arg Phe Glu Thr Glu Leu Phe Arg Gln Phe Arg
50              55              60
Val Gln Ser Val Ser Pro Ala Arg Val Ala Ser Pro Pro Met Lys Leu
65              70              75              80
Pro Gly Arg Phe Thr Ser Gly Leu Ile Leu Leu Phe Thr Ser Cys Gly
              85              90              95
Ala Leu Ala Gln Ser Glu Leu Asp Val Arg Ile Lys Pro Ser Asn Asp
              100              105              110
Ala

```

<210> 141
 <211> 324

<212> DNA

<213> Homo sapiens

<400> 141

gaattcctct tggatagctt cgggtaaatg ggtacagcaa atatcaggag cgcaaccgca
 60
 acctttactt actggtacat gaacaccatt tacattacag ctatcgtagt caccaccagt
 120
 catgtgaaca gacacataac tgaaagggtt ataaaccaca gtctcacggt acgtatgacc
 180
 gtcaactgtg aacaccgcta agtaatagcc tgcgggggct tgcatagaact cctttgacca
 240
 tgcgtaataa atacgtccgt cattagtcac acctgatggg gcgaaacaaa aagaacggca
 300
 gcagttatca ccgcccatat gcgt
 324

<210> 142

<211> 106

<212> PRT

<213> Homo sapiens

<400> 142

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Gly | Gly | Asp | Asn | Cys | Cys | Arg | Ser | Phe | Cys | Phe | Ala | Pro | Ser | Gly |
| 1 | | | | 5 | | | | 10 | | | | | 15 | | |
| Val | Thr | Asn | Asp | Gly | Arg | Ile | Tyr | Tyr | Ala | Trp | Ser | Lys | Glu | Phe | Met |
| | | 20 | | | | | 25 | | | | | 30 | | | |
| Gln | Ala | Pro | Ala | Gly | Tyr | Tyr | Leu | Ala | Val | Phe | Thr | Val | Asp | Gly | His |
| | 35 | | | | | 40 | | | | | 45 | | | | |
| Thr | Tyr | Arg | Glu | Thr | Val | Val | Tyr | Lys | Pro | Phe | Ser | Tyr | Val | Ser | Val |
| | 50 | | | | | 55 | | | | | 60 | | | | |
| His | Met | Thr | Trp | Gly | Glu | Tyr | Asp | Ser | Cys | Asn | Val | Asn | Gly | Val | His |
| 65 | | | | 70 | | | | | 75 | | | | 80 | | |
| Val | Pro | Val | Ser | Lys | Gly | Cys | Gly | Cys | Ala | Pro | Asp | Ile | Cys | Cys | Thr |
| | | | 85 | | | | | 90 | | | | | 95 | | |
| His | Leu | Pro | Glu | Ala | Ile | Gln | Glu | Glu | Phe | | | | | | |
| | | | 100 | | | | | 105 | | | | | | | |

<210> 143

<211> 1325

<212> DNA

<213> Homo sapiens

<400> 143

nacgcgtgga tctgccagct gagcctggag ctgtgcaggc agctgccctg ctacgatgag
 60
 gcaccccgagg agaagaactt cctgtacaaa tgcataggca ccaccctggg tgctgcttca
 120
 agtaaggagg tggtagaggaa gcaccttcaa gagctgctgg agacggccag ataccaggag
 180
 gaggcagaac gcgagggcct cgctgctgc ttcgggatct gtgccatctc ccacctcgag
 240
 gacacgctgg ccagctgga ggacttcgtg aggtcagagg tcttcagaaa atccattggc
 300

attctcaaca tttttaagga tcgaagtgag aacgaagtgg agaaggtgaa gagtgtctctg
 360
 atcctgtgct atgggcacgt ggcggcccg gccccccggg agctgggtgct ggccaaggta
 420
 gagtgcagaca tcctccggaa catcntgcc gcaattcagc acnncaagga cccagccctg
 480
 aagctgtgcc ttgtccagag tgtgtgcatg gtcagccgcg ccatctgcag cagcaccag
 540
 gctggctcct tccacttcac ccggaaagca gagctggtgg cacagatgat ggagttcatc
 600
 agggcagagc ccccgactc cttgaggaca cctattcgga agaaagccat gctcacctgc
 660
 acttacttgg tctccgtgga gccagcgctg gacgagcagg cccgggcgga tgtgatccat
 720
 ggctgcctgc acagcatcat ggccctgctg cctgagccca aggaggagga cggaggctgc
 780
 cagaagtccc tgtatctgga gacactgcac gcccttgagg atctgctgac gagcctcctg
 840
 cagcggaaca tgacccccca aggctgcag atcatgattg agcacctgag cccatggatc
 900
 aagtccccaa gaggtcacgt agcggcgctg gccctaggcc tgagcgccct cctcgtgcgc
 960
 tacttcctgg agcacctgcg tgtcagtggc gccaagtag ataccaggtt tccatctgag
 1020
 cccaggatcc tgtgcaatgg ccctgggtgcc cttccacaac ctgggccttc tcacggcct
 1080
 cttctcccca cgggtgtgcgg acctgtggcc tgccaccgc caggaggccg tggactgtgt
 1140
 ctactccctg ctgtacctcc agctcggcta tgagggttc tcccgggact accgcgatga
 1200
 cgtggcggag cggctcctca gcctcaagga cggcctcgtg caccctgacc ccgccattct
 1260
 cttccacacc tgccacagtg taggccagat tattgccaag cgcctcccc cagcccttca
 1320
 cgcgt
 1325

<210> 144

<211> 390

<212> PRT

<213> Homo sapiens

<400> 144

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Xaa | Ala | Trp | Ile | Cys | Gln | Leu | Ser | Leu | Glu | Leu | Cys | Arg | Gln | Leu | Pro |
| 1 | | | 5 | | | | | 10 | | | | | 15 | | |
| Cys | Tyr | Asp | Glu | Ala | Pro | Gln | Glu | Lys | Asn | Phe | Leu | Tyr | Lys | Cys | Ile |
| | | 20 | | | | | | 25 | | | | 30 | | | |
| Gly | Thr | Thr | Leu | Gly | Ala | Ala | Ser | Ser | Lys | Glu | Val | Val | Arg | Lys | His |
| | | 35 | | | | | 40 | | | | 45 | | | | |
| Leu | Gln | Glu | Leu | Leu | Glu | Thr | Ala | Arg | Tyr | Gln | Glu | Glu | Ala | Glu | Arg |
| | | 50 | | | | 55 | | | | 60 | | | | | |
| Glu | Gly | Leu | Ala | Cys | Cys | Phe | Gly | Ile | Cys | Ala | Ile | Ser | His | Leu | Glu |
| 65 | | | | 70 | | | | 75 | | | 80 | | | | |
| Asp | Thr | Leu | Ala | Gln | Leu | Glu | Asp | Phe | Val | Arg | Ser | Glu | Val | Phe | Arg |

| 85 | | | | | | | | | | 90 | | | | | 95 | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|--|--|--|--|
| Lys | Ser | Ile | Gly | Ile | Leu | Asn | Ile | Phe | Lys | Asp | Arg | Ser | Glu | Asn | Glu | | | | |
| 100 | | | | | | | | | | 105 | | | | | 110 | | | | |
| Val | Glu | Lys | Val | Lys | Ser | Ala | Leu | Ile | Leu | Cys | Tyr | Gly | His | Val | Ala | | | | |
| 115 | | | | | | | | | | 120 | | | | | 125 | | | | |
| Ala | Arg | Ala | Pro | Arg | Glu | Leu | Val | Leu | Ala | Lys | Val | Glu | Ser | Asp | Ile | | | | |
| 130 | | | | | | | | | | 135 | | | | | 140 | | | | |
| Leu | Arg | Asn | Ile | Xaa | Pro | Ala | Leu | Gln | His | Xaa | Lys | Asp | Pro | Ala | Leu | | | | |
| 145 | | | | | | | | | | 150 | | | | | 155 | | | | |
| Lys | Leu | Cys | Leu | Val | Gln | Ser | Val | Cys | Met | Val | Ser | Arg | Ala | Ile | Cys | | | | |
| 165 | | | | | | | | | | 170 | | | | | 175 | | | | |
| Ser | Ser | Thr | Gln | Ala | Gly | Ser | Phe | His | Phe | Thr | Arg | Lys | Ala | Glu | Leu | | | | |
| 180 | | | | | | | | | | 185 | | | | | 190 | | | | |
| Val | Ala | Gln | Met | Met | Glu | Phe | Ile | Arg | Ala | Glu | Pro | Pro | Asp | Ser | Leu | | | | |
| 195 | | | | | | | | | | 200 | | | | | 205 | | | | |
| Arg | Thr | Pro | Ile | Arg | Lys | Lys | Ala | Met | Leu | Thr | Cys | Thr | Tyr | Leu | Val | | | | |
| 210 | | | | | | | | | | 215 | | | | | 220 | | | | |
| Ser | Val | Glu | Pro | Ala | Leu | Asp | Glu | Gln | Ala | Arg | Ala | Asp | Val | Ile | His | | | | |
| 225 | | | | | | | | | | 230 | | | | | 235 | | | | |
| Gly | Cys | Leu | His | Ser | Ile | Met | Ala | Leu | Leu | Pro | Glu | Pro | Lys | Glu | Glu | | | | |
| 245 | | | | | | | | | | 250 | | | | | 255 | | | | |
| Asp | Gly | Gly | Cys | Gln | Lys | Ser | Leu | Tyr | Leu | Glu | Thr | Leu | His | Ala | Leu | | | | |
| 260 | | | | | | | | | | 265 | | | | | 270 | | | | |
| Glu | Asp | Leu | Leu | Thr | Ser | Leu | Leu | Gln | Arg | Asn | Met | Thr | Pro | Gln | Gly | | | | |
| 275 | | | | | | | | | | 280 | | | | | 285 | | | | |
| Leu | Gln | Ile | Met | Ile | Glu | His | Leu | Ser | Pro | Trp | Ile | Lys | Ser | Pro | Arg | | | | |
| 290 | | | | | | | | | | 295 | | | | | 300 | | | | |
| Gly | His | Val | Ala | Ala | Arg | Ala | Leu | Gly | Leu | Ser | Ala | Leu | Leu | Val | Arg | | | | |
| 305 | | | | | | | | | | 310 | | | | | 315 | | | | |
| Tyr | Phe | Leu | Glu | His | Leu | Arg | Val | Ser | Gly | Ala | Gln | Val | Asp | Thr | Arg | | | | |
| 325 | | | | | | | | | | 330 | | | | | 335 | | | | |
| Phe | Pro | Ser | Glu | Pro | Arg | Ile | Leu | Cys | Asn | Gly | Pro | Gly | Ala | Leu | Pro | | | | |
| 340 | | | | | | | | | | 345 | | | | | 350 | | | | |
| Gln | Pro | Gly | Pro | Ser | His | Arg | Pro | Leu | Leu | Pro | Thr | Val | Cys | Gly | Pro | | | | |
| 355 | | | | | | | | | | 360 | | | | | 365 | | | | |
| Val | Ala | Cys | His | Pro | Pro | Gly | Gly | Arg | Gly | Leu | Cys | Leu | Leu | Pro | Ala | | | | |
| 370 | | | | | | | | | | 375 | | | | | 380 | | | | |
| Val | Pro | Pro | Ala | Arg | Leu | | | | | | | | | | | | | | |
| 385 | | | | | | | | | | | | | | | 390 | | | | |

<210> 145

<211> 802

<212> DNA

<213> Homo sapiens

<400> 145

400-115
cggccgtcta ggtccggctc agtgcgctgt tgctcgccgt agaacacgag gctgcgcaag
60 cataagcaga cgtagagagt ggtcacatcc atgtcgatgg tgtgcgcgta atgaaggtct
120 acatcaccct ggtgaaggcc tgcaccacta gcgtcggcac catttccccg cgtcggacaa
180 gacatcatgc cccatatctt gacagaatgt ctgacatgag tatgccacgc cgagcagcac
240

cagaggacga caccgatctg gcggacgccg cccgttcacg gcgcagatac ctcacccctg
 300
 tcatttggg cgttatcgtc gctgtcctcg gactaggcat tttcgggtat cttgcgtggg
 360
 ggtcattgtg cgatcaagct gccgggggtct gtcagcgtgg tgaacccggt atgtactggg
 420
 gttcgggtgg ctctctggcc attctcggac tcattatcgg ggtcttgacg cagatctggc
 480
 tggagaagcg ctggtggcac atgcttgcca tcgtcatccc ggctgttttc atcgtcgccg
 540
 gtatcttttt ctggctcgcc gtctaagaag gggcgtcaca gattccacaa acgacacagg
 600
 tattgatctc cgttttatcg gctcctagca gccgtgggtca acgtatcgct atcaagcgat
 660
 acaggactcg tcgttcgcat cgttggtgtg ctgctgggaa acaatcccag cgatctactc
 720
 ggctaccgcc agacagttca ctcacaaccc ctcacgccgg cgcagacatc aaatcccatt
 780
 ctcgatagac ggccacacc ac
 802

<210> 146

<211> 151

<212> PRT

<213> Homo sapiens

<400> 146

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Lys | Val | Tyr | Ile | Thr | Leu | Val | Lys | Ala | Cys | Thr | Thr | Ser | Val | Gly |
| 1 | | | | 5 | | | | | 10 | | | | | 15 | |
| Thr | Ile | Ser | Pro | Arg | Arg | Thr | Arg | His | His | Ala | Pro | Tyr | Leu | Asp | Arg |
| | | | 20 | | | | | 25 | | | | | 30 | | |
| Met | Ser | Asp | Met | Ser | Met | Pro | Arg | Arg | Ala | Ala | Pro | Glu | Asp | Asp | Thr |
| | | 35 | | | | 40 | | | | | | 45 | | | |
| Asp | Leu | Ala | Asp | Ala | Ala | Arg | Ser | Trp | Arg | Arg | Tyr | Leu | Ile | Leu | Val |
| | 50 | | | | | 55 | | | | | 60 | | | | |
| Ile | Cys | Gly | Val | Ile | Val | Ala | Val | Leu | Gly | Leu | Gly | Ile | Phe | Gly | Tyr |
| 65 | | | | 70 | | | | 75 | | | | | 80 | | |
| Leu | Ala | Trp | Trp | Ser | Leu | Cys | Asp | Gln | Ala | Ala | Gly | Val | Cys | Gln | Arg |
| | | | 85 | | | | | 90 | | | | | 95 | | |
| Gly | Glu | Pro | Val | Met | Tyr | Trp | Cys | Ser | Val | Val | Ser | Leu | Ala | Ile | Leu |
| | | | 100 | | | | | 105 | | | | | 110 | | |
| Gly | Leu | Ile | Ile | Gly | Val | Leu | Thr | Gln | Ile | Trp | Leu | Glu | Lys | Arg | Trp |
| | 115 | | | | | 120 | | | | | | 125 | | | |
| Trp | His | Met | Leu | Ala | Ile | Val | Ile | Pro | Ala | Val | Phe | Ile | Val | Ala | Gly |
| | 130 | | | | | 135 | | | | | 140 | | | | |
| Ile | Phe | Phe | Trp | Leu | Ala | Val | | | | | | | | | |
| 145 | | | | | | 150 | | | | | | | | | |

<210> 147

<211> 368

<212> DNA

<213> Homo sapiens

<400> 147

acgcgtgaaa acggtatgac tcttctggcc ttagtagatc tgtctaaaaa acccgatgag
 60
 ttacacagt gggcattagt agcccgcgat gttcatgaca ttcctgggtc acgaaaagtt
 120
 attggtcaga aagtaccttg tggtgcagtg acgggggtcgg aaaagggtgct tcataaaaaag
 180
 gattactggg atctagcaac acctatgccca attgcgtggg gtacaacgga ccgaacagtt
 240
 attgctgatg cacgacgtac aatccccacc acggagtggg atatccttgc aagactacgt
 300
 ccacgcctag aagagggttcg caagcaacgt aatgatgtat tgctcctcaa cgaggaggat
 360
 ccccccta
 368

<210> 148

<211> 117

<212> PRT

<213> Homo sapiens

<400> 148

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Thr | Leu | Leu | Ala | Leu | Val | Asp | Leu | Ser | Lys | Lys | Pro | Asp | Glu | Phe |
| 1 | | | | 5 | | | | | 10 | | | | | 15 | |
| Thr | Gln | Trp | Ala | Leu | Val | Ala | Arg | Asp | Val | His | Asp | Ile | Pro | Gly | Leu |
| | | | 20 | | | | | 25 | | | | | 30 | | |
| Arg | Lys | Val | Ile | Gly | Gln | Lys | Val | Pro | Cys | Val | Ala | Val | Thr | Gly | Ser |
| | | 35 | | | | 40 | | | | | | 45 | | | |
| Glu | Lys | Val | Leu | His | Lys | Lys | Asp | Tyr | Trp | Asp | Leu | Ala | Thr | Pro | Met |
| | 50 | | | | 55 | | | | | | 60 | | | | |
| Pro | Ile | Ala | Trp | Gly | Thr | Thr | Asp | Arg | Thr | Val | Ile | Ala | Asp | Ala | Arg |
| 65 | | | | 70 | | | | 75 | | | | | 80 | | |
| Arg | Thr | Ile | Pro | Thr | Thr | Glu | Trp | Asp | Ile | Leu | Ala | Arg | Leu | Arg | Pro |
| | | | 85 | | | | | 90 | | | | | 95 | | |
| Arg | Leu | Glu | Glu | Val | Arg | Lys | Gln | Arg | Asn | Asp | Val | Leu | Leu | Leu | Asn |
| | | | 100 | | | | 105 | | | | | | 110 | | |
| Glu | Glu | Asp | Pro | Pro | | | | | | | | | | | |
| | | | 115 | | | | | | | | | | | | |

<210> 149

<211> 407

<212> DNA

<213> Homo sapiens

<400> 149

nngctagcat ggaccctagt cacacaggca gccatacccg aggtcaaagt gacccatttt
 60
 cctaataatgg ccgctcagat ccaatacttt gaagattcgt ccgtgggttat atggcacgat
 120
 gcgggtggatg gstatcgtgta ccgaagtgcg gatgaaggca agtcgtgggc cccaattaag
 180
 gggcctgaac agggtcaggc gcaccttttc gtgctccatc cctacgacaa gactcaagcg
 240
 tatattctga cgcgcagcac tcagcattgg cgcacgtcga accgtggcga gacgtggcag
 300

tcattctcaa cgctcatcc gcctacgacc ttgaaagcta tgctcttgga ctttcatccg
 360
 acgcatcatg actggatcct tttcacgggc caggcttgca cggtaaa
 407

<210> 150
 <211> 135
 <212> PRT
 <213> Homo sapiens

<400> 150
 Xaa Leu Ala Trp Thr Leu Val Thr Gln Ala Ala Ile Pro Glu Val Lys
 1 5 10 15
 Val Thr His Phe Pro Asn Met Ala Ala Gln Ile Gln Tyr Phe Glu Asp
 20 25 30
 Ser Ser Val Val Ile Trp His Asp Ala Val Asp Gly Ile Val Tyr Arg
 35 40 45
 Ser Ala Asp Glu Gly Lys Ser Trp Ala Pro Ile Lys Gly Pro Glu Gln
 50 55 60
 Gly Gln Ala His Leu Phe Val Leu His Pro Tyr Asp Lys Thr Gln Ala
 65 70 75 80
 Tyr Ile Leu Thr Arg Ser Thr Gln His Trp Arg Thr Ser Asn Arg Gly
 85 90 95
 Glu Thr Trp Gln Ser Phe Ser Thr Pro His Pro Pro Thr Thr Leu Lys
 100 105 110
 Ala Met Pro Leu Asp Phe His Pro Thr His His Asp Trp Ile Leu Phe
 115 120 125
 Thr Gly Gln Ala Cys Thr Val
 130 135

<210> 151
 <211> 448
 <212> DNA
 <213> Homo sapiens

<400> 151
 accggtgtcc gtggctattg ccccgaaatgg tccccatccg cgccccggg aactcctcg
 60
 gcttttcgcg catccaggtc cccagcccca gctactggtg cgccccgagc ccctaggtgc
 120
 cagagcgggtg gtcggccggg ctccctgcca gtctcggtc ctccctctc cccaccagaa
 180
 ggaaaaactt gggcccttcg agaaccctgt ggaatgttct ttgtaatcaa ctgtacatcc
 240
 gcttcacgg cacggcctcg tgcaaaatcg cgggtttcgg ggccttgag caaattgcgc
 300
 ttgtcagcgg cgacgtcagg aggacaagg gaggggttcg cggctgaaac tgcagcttcg
 360
 cagcacagag ccattttagg ctgctcccca cctcgcgggg cccatgggaa gccggccccg
 420
 ggagggcgcg gctgcatgga tattcgac
 448

<210> 152

<211> 149
 <212> PRT
 <213> Homo sapiens

<400> 152
 Thr Gly Val Arg Gly Tyr Cys Pro Glu Trp Ser Pro Ser Ala Ser Pro
 1 5 10 15
 Gly Thr Pro Ser Ala Phe Arg Ala Ser Arg Ser Pro Ala Pro Ala Thr
 20 25 30
 Gly Ala Pro Arg Ala Pro Arg Cys Gln Ser Gly Gly Arg Pro Gly Ser
 35 40 45
 Cys Pro Val Ser Ala Pro Pro Ser Ser Pro Pro Glu Gly Lys Thr Trp
 50 55 60
 Ala Leu Arg Glu Pro Cys Gly Met Phe Phe Val Ile Asn Cys Thr Ser
 65 70 75 80
 Ala Ser Thr Ala Arg Pro Arg Ala Lys Ser Arg Val Ser Gly Pro Trp
 85 90 95
 Ser Lys Leu Arg Leu Ser Ala Ala Thr Ser Gly Gly Gln Gly Glu Gly
 100 105 110
 Phe Ala Ala Glu Thr Ala Ala Ser Gln His Arg Ala Ile Leu Gly Cys
 115 120 125
 Ser Pro Pro Arg Gly Ala His Gly Lys Pro Ala Pro Gly Gly Arg Gly
 130 135 140
 Cys Met Asp Ile Arg
 145

<210> 153
 <211> 440
 <212> DNA
 <213> Homo sapiens

<400> 153
 nnntgggtcc atgtatgtgt gtgtatatga gggagacacg caggtgtgtg tccgagtgtg
 60
 tgtccatggg tccatgtatg tgtgtgtata tgtgggggaa caggtgtgtg tccgagtgtg
 120
 tgcattgggtc cgtgtatatg cgtgtatatata tgcgggggata tgtatatgtg tgtgtgtatg
 180
 aacaggtgta agtgggggagc actcaggtgt gtctgtgtgt gttcgtgtac acgtgtgtaa
 240
 gtgggtgacc atgaaggggt gtgtgtgtcc gtgtgtaggt ttgcgtgcat gcacacatgc
 300
 atgtgtgtac tggggcatcc aagcccctgg tctccactcc attccaccct acgcctacct
 360
 ccttgatctc tgcgcccagc cttggctgtg ctcccctgct gtagtcacgt ggggtgtctgc
 420
 acgtgggtgt ctgcacgcgt
 440

<210> 154
 <211> 69
 <212> PRT
 <213> Homo sapiens

<400> 154

Gly Arg His Ala Gly Val Cys Pro Ser Val Cys Pro Trp Val His Val
 1 5 10 15
 Cys Val Cys Ile Cys Gly Gly Thr Gly Val Cys Pro Ser Val Cys Met
 20 25 30
 Gly Pro Cys Ile Cys Val Tyr Ile Cys Gly Asp Met Tyr Met Cys Val
 35 40 45
 Cys Met Asn Arg Cys Lys Trp Gly Ala Leu Arg Cys Val Cys Val Cys
 50 55 60
 Ser Cys Thr Arg Val
 65

<210> 155

<211> 344

<212> DNA

<213> Homo sapiens

<400> 155

acgcgtatcg accaccatgt cgtcgtcacc acggcaagcg ctctcggcgg gcgagaacga
 60
 gtgaacatgg ccgagttgat ggccgatgcc gcgaccggca cgaaaccgtc ctacctacag
 120
 cgatcttctt cctcgatcac ctcgtttgaa gtggacaggg aacaaagaca ctgagacaac
 180
 gcgccgcagg aagtaaaaag ttcgctctcc gatcacggcc gtcgcgcgag tgcacagggg
 240
 gaactgggca cctcgcaagc tacgccaccg cgatccatgc ccccgcccgt atcttcgcgc
 300
 tcctctacct ccccttacc gatcagcatt atatccgac taga
 344

<210> 156

<211> 92

<212> PRT

<213> Homo sapiens

<400> 156

Met Ala Glu Leu Met Ala Asp Ala Ala Thr Gly Thr Lys Pro Ser Tyr
 1 5 10 15
 Leu Gln Arg Ser Ser Ser Ile Thr Ser Phe Glu Val Asp Arg Glu
 20 25 30
 Gln Arg His Ser Asp Asn Ala Pro Gln Glu Val Lys Ser Ser Leu Ser
 35 40 45
 Asp His Gly Arg Arg Ala Ser Ala Gln Gly Glu Leu Gly Thr Ser Gln
 50 55 60
 Ala Thr Pro Pro Arg Ser Met Pro Pro Pro Val Ser Ser Ala Ser Ser
 65 70 75 80
 Thr Ser Pro Leu Pro Ile Ser Ile Ile Ser Asp Leu
 85 90

<210> 157

<211> 6816

<212> DNA

<213> Homo sapiens

<400> 157
nnagatctcc aaagaaccaa aaggatcaat atttctggat tcctgtatgg gtgtcgttca
60
gaacaacaaa gtcaggcgtt ttgcttttga gctcaagatg caggacaaaa gtagttatct
120
cttggcagca gacagtgaag tggaaatgga agaattggatc acaattctaa ataagatcct
180
ccagctcaac tttgaagctg caatgcaaga aaagcgaaat ggcgactctc acgaagatga
240
tgaacaaagc aaattggaag gttctggttc cggtttagat agctacctgc cggaacttgc
300
caagagtgc agagaagcag aaatcaaact gaaaagtga agcagagtca aactttttta
360
tttggaccca gatgccaga agcttgactt ctcatcagct gagccagaag tgaagtcatt
420
tgaagagaag tttggaaaaa ggatccttgt caagtgcatt gatttatctt tcaatttgca
480
atgctgtgtt gccgaaaatg aagaaggacc cactacaaat gttgaacctt tctttgttac
540
tctatccctg tttgacataa aatacaaccg gaagatttct gccgatttcc acgtagacct
600
gaaccatttc tcagtggaggc aaatgatcgc caccacgtcc ccggcgctga tgaatggcag
660
tgggccgaaa cccaatctgc cctcaggggc atccttcacg aagccgccat gcagtatccg
720
aagcagggaa tattttcagt cacttgctct catccagata tatttcttgt ggccagaatt
780
gaaaaagtcc ttcaggggag catcacacat tgcgctgagc catatatgaa aagttcagac
840
tcttctaagg tggccagaa ggtgctgaag aatgccaaag aggcattgcca aagactagga
900
cagtatagaa tgccatttgc ttgggcagca aggacattgt ttaaggatgc atctggaaat
960
cttgacaaaa atgccagatt ttctgccatc tacaggcaag acagcaataa gctatccaat
1020
gatgacatgc tcaagttact tgcagacttt cggaaacctg agaagatggc taagctccca
1080
gtgatttttag gcaatctaga cattacaatt gataatgttt cctcagactt ccctaattat
1140
gttaattcat catacattcc cacaaaaaa tttgaaacct gcagtaaaac tcccatcacg
1200
tttgaagtgg aggaatttgt gccctgcata ccaaaacaca ctcagcctta caccatctac
1260
accaatcacc tttacgttta tcctaagtac ttgaaatacg acagtcagaa gtcttttgcc
1320
aaggctagaa atattgcat ttgcattgaa ttcaaagatt cagatgagga agactctcag
1380
ccccctaagt gcatttatgg cagacctggt gggccagttt tcacaagaag cgcctttgct
1440
gcagttttac accatcacca aaaccagaa ttttatgatg agattaaaat agagttgccc
1500
actcagctgc atgaaaagca ccacctgttg ctacattct tccatgtcag ctgtgacaac
1560

tcaagtaaag gaagcacgaa gaagagggat gtcgttgaaa cccaagttgg ctactcctgg
1620
cttccccctcc tgaaagacgg aaggggtggtg acaagcgagc agcacatccc ggtctcggcg
1680
aaccttcctt cgggctatct tggctaccag gagcttgga tgggcaggca ttatgggtccg
1740
gaaattaaat gggtagatgg aggcaagcca ctgctgaaaa tttccactca tctgggtttct
1800
acagtgtata ctcaggatca gcatttacat aatTTTTTcc agtactgtca gaaaaccgaa
1860
tctggagccc aagccttagg aaacgaactt gtaaagtacc ttaagagtct gcatgcgatg
1920
gaagggccacg tgatgatcgc cttcttgccc actatcctaa accagctgtt ccgagtcctc
1980
accagagcca cacaggaaga agtcgcggtt aacgtgactc gggtcattat tcatgtggtt
2040
gcccagtgcc atgaggaagg attggagagc cacttgaggt catatgttaa gtacgcgtat
2100
aaggctgagc catatgttgc ctctgaatac aagacagtgc atgaagaact gaccaaactc
2160
atgaccacga ttctcaagcc ttctgccgat ttcctcacca gcaacaaact actgaagtac
2220
tcatgggttt tctttgatgt actgatcaaa tctatggctc agcatttgat agagaactcc
2280
aaagttaagt tgctgcgaaa ccagagattt cctgcacctc atcatcatgc agtggaaacc
2340
gttgtaaata tgctgatgcc acacatcact cagaagtttc gagataatcc agaggcatct
2400
aagaacgcga atcatagcct tgctgtcttc atcaagagat gtttcacctt catggacagg
2460
ggctttgtct tcaagcagat caacaactac attagctgtt ttgctcctgg agacccaaag
2520
accctctttg aatacaagtt tgaatttctc cgtgtagtgt gcaaccatga acattatatt
2580
ccgttgaaact taccaatgcc atttggaaaa ggcaggattc aaagatacca agacctccag
2640
cttgactact cattaacaga tgagttctgc agaaaccact tcttggtggg actgttactg
2700
agggaggtgg ggacagccct ccaggagtgc cgggaggtcc gtctgatcgc catcagtgtg
2760
ctcaagaacc tgctgataaa gcattctttt gatgacagat atgcttcaag gagccatgag
2820
gcaaggatag ccacctcta cctgcctctg tttggtctgc tgattgaaaa cgtccagcgg
2880
atcaatgtga gggatgtgtc acccttcctt gtgaacgcgg gcatgactgt gaaggatgaa
2940
tccctggctc taccagctgt gaatccgtg gtgacgccgc agaagggaag caccctggac
3000
aacagcctgc acaaggacct gctgggcgcc atctccggca ttgcttctcc atatacaacc
3060
tcaactccaa acatcaacag tgtgagaaat gctgattcga gaggatctct cataagcaca
3120
gattcgggta acagccttcc agaaaggaat agtgagaaga gcaattcctt ggataagcac
3180

caacaaagta gcacattggg aaattccgtg gttcgtctg ataaacttga ccagtctgag
3240
attaagagcc tactgatgtg tttcctctac atcttaaaga gcatgtctga tgatgctttg
3300
tttcatatt ggaacaaggc ttcaacatct gaacttatgg atttttttac aatatctgaa
3360
gtctgcctgc accagttcca gtacatgggg aagcgataca tagccagaac aggaatgatg
3420
catgccagat tgcagcagct gggcagcctg gataactctc tcaactttta ccacagctat
3480
ggccactcgg acgcagatgt tctgcaccag tcattacttg aagccaacat tgctactgag
3540
gtttgcctga cagctctgga cacgctttct ctatttacat tggcgtttta gaaccagctc
3600
ctggccgacc atggacataa tcctctcatg aaaaaagttt ttgatgtcta cctgtgtttt
3660
cttcaaaaac atcagtctga aacggcttta aaaaatgtct tcaactgcctt aaggctctta
3720
atttataagt ttccctcaac attctatgaa gggagagcgg acatgtgtgc ggctctgtgt
3780
tacgagattc tcaagtgtg taactccaag ctgagctcca tcaggacgga ggcctcccag
3840
ctgctctact tcctgatgag gaacaacttt gattacactg gaaagaagtc ctttgtccgg
3900
acacatttgc aagtcatcat atctgtcagc cagctgatag cagacgttgt tggcattggg
3960
ggaaccagat tccagcagtc cctgtccatc atcaacaact gtgccaacag tgaccggctt
4020
attaagcaca ccagcttctc ctctgatgtg aaggacttaa ccaaaggat acgcacgggtg
4080
ctaattggcca ccgcccagat gaaggagcat gagaacgacc cagagatgct ggtggacctc
4140
cagtacagcc tggccaaatc ctatgccagc acgcccagac tcaggaagac gtggctcgac
4200
agcatggcca ggatccatgt caaaaatggc gatctctcag aggcagcaat gtgctatgtc
4260
cacgtaacag ccctagtggc agaatatctc acacggaaag aagcagtcca gtgggagccg
4320
ccccttctcc cccacagcca tagcgctgc ctgaggagga gccggggagg cgtgtttaga
4380
caaggatgca ccgccttcag ggtcattacc ccaaaccatc acgaggaggc ctccatgatg
4440
gaagacgtgg ggatgcagga tgtccatttc aacgaggatg tgctgatgga gctccttgag
4500
cagtgcgag atggactctg gaaagccgag cgctacgagc tcattgccga catctacaaa
4560
cttatcatcc ccatttatga gaagcggagg gatthttgaga ggctggccca tctgtatgac
4620
acgtgcacc gggcctacag caaagtgacc gaggtcatgc actcgggccc caggcttctg
4680
gggacctact tccgggtagc cttcttcggg caggcagcgc aataccagtt tacagacagt
4740
gaaacagatg tggagggatt ctttgaagat gaagatggaa aggagtatat ttacaaggaa
4800

cccaaaactca caccgctgtc ggaaatttct cagagactcc ttaaactgta ctcggataaa
4860
tttggttctg aaaatgtcaa aatgatacag gattctggca aggtcaaccc taaggatctg
4920
gattctaagt atgcctacat ccagggtgact cacgtcatcc ccttctttga cgaaaaagag
4980
ttgcaagaaa ggaaaacaga gtttgagaga tcccacaaca tccgccgctt catgtttgag
5040
atgccattta cgcagaccgg gaagaggcag ggcgggggtg aagagcagtg caaacggcgc
5100
accatcctga cagccataca ctgcttcctt tatgtgaaga agcgcacccc tgtcatgtac
5160
cagcaccaca ctgacctgaa ccccatcgag gtggccattg acgagatgag taagaagggtg
5220
gcggagctcc ggagctgtg ctctcgggc gaggtggaca tgatcaaact gcagctcaaa
5280
ctccagggca gcgtgagtgt tcagggtcaat gctggccac tagcatatgc gcgagctttc
5340
ttagatgata caaacacaaa gcgatatcct gacaataaag tgaagctgct taagggaagtt
5400
ttcaggcaat ttgtggaagc ttgcggtcaa gccttagcgg taaacgaacg tctgattaaa
5460
gaagaccagc tcgagtatca ggaagaaatg aaagccaact acagggaaat ggcaaggag
5520
ctttctgaaa tcatgcatga gcagatctgc cccctggagg agaagacgag cgtcttaccg
5580
aattcccttc acatcttcaa cgccatcagt gggactccaa caagcacaat gggtcacggg
5640
atgaccagct cgtcttcggg cgtgtgatta catctcatgg cccgtgtgtg gggacttgct
5700
ttgtcatttg caaactcagg atgctttcca aagccaatca ctggggagac cgagcacagg
5760
gaggaccaag ggaaggggga gagaaaggaa ataaagaaca acgttatattc ttaacagact
5820
ttctatagga gttgtaagaa ggtgcacata tttttttaa tctcactggc aatattcaaa
5880
gttttcattg tgtcttaaca aagggtgtgt agacactctt gagctggact tagattttat
5940
tcttccttgc agagtagtgt tagaatagat ggcctacaga aaaaaaagggt tctgggatct
6000
acatggcagg gagggctgca ctgacattga tgcctggggg accttttgcc tcgaggctga
6060
gctggaaaat cttgaaaata tttttttttt cctgtggcac attcagggtg aatacaagaa
6120
ctatttttgt gactagtttt tgatgacctt agggaaactga ccattgtaat tttgtacca
6180
gtgaaccagg agatttagtg cttttatatt catttccttg catttaagaa aatatgaaag
6240
cttaaggaat tatgtgagct taaaactagt caagcagttt agaaccaaag gcctatatta
6300
ataaccgcaa ctatgctgaa aagtacaaag tagtacagta tattgttatg tacatatcat
6360
tgttaataca gtcctggcat tctgtacata tatgtattac atttctacat ttttaatact
6420

cacatgggct tatgcattaa gtttaattgt gataaatttg tgctgttcca gtatatgcaa
 6480
 tacacttttaa tgttttattc ttgtacataa aaatgtgcaa tatggagatg tatacagtct
 6540
 ttactatatt aggtttataa acagttttta gaatttcac cttttgccaa aatgggtggag
 6600
 tatgtaattg gtaaatcata aatcctgtgg tgaatgggtg tgtactttta agctgtcacc
 6660
 atgttatatt ttcttttaag acattaattt agtaatttta tatttgggaa aataaaggtt
 6720
 tttaatttta ttttaactgga atcactgccc tgctgtaatt aaacattctg taccacatct
 6780
 gtattaaaaa gacattgctg accattaaaa aaaaaa
 6816

<210> 158

<211> 1572

<212> PRT

<213> Homo sapiens

<400> 158

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Ala | Ser | Gly | Asn | Leu | Asp | Lys | Asn | Ala | Arg | Phe | Ser | Ala | Ile | Tyr | Arg |
| 1 | | | 5 | | | | | | 10 | | | | | 15 | |
| Gln | Asp | Ser | Asn | Lys | Leu | Ser | Asn | Asp | Asp | Met | Leu | Lys | Leu | Leu | Ala |
| | | | 20 | | | | | 25 | | | | | 30 | | |
| Asp | Phe | Arg | Lys | Pro | Glu | Lys | Met | Ala | Lys | Leu | Pro | Val | Ile | Leu | Gly |
| | | 35 | | | | | 40 | | | | | 45 | | | |
| Asn | Leu | Asp | Ile | Thr | Ile | Asp | Asn | Val | Ser | Ser | Asp | Phe | Pro | Asn | Tyr |
| | 50 | | | | | 55 | | | | | 60 | | | | |
| Val | Asn | Ser | Ser | Tyr | Ile | Pro | Thr | Lys | Gln | Phe | Glu | Thr | Cys | Ser | Lys |
| 65 | | | | | 70 | | | | | 75 | | | | 80 | |
| Thr | Pro | Ile | Thr | Phe | Glu | Val | Glu | Glu | Phe | Val | Pro | Cys | Ile | Pro | Lys |
| | | | 85 | | | | | | 90 | | | | | 95 | |
| His | Thr | Gln | Pro | Tyr | Thr | Ile | Tyr | Thr | Asn | His | Leu | Tyr | Val | Tyr | Pro |
| | | | 100 | | | | | 105 | | | | | 110 | | |
| Lys | Tyr | Leu | Lys | Tyr | Asp | Ser | Gln | Lys | Ser | Phe | Ala | Lys | Ala | Arg | Asn |
| | | 115 | | | | | 120 | | | | | 125 | | | |
| Ile | Ala | Ile | Cys | Ile | Glu | Phe | Lys | Asp | Ser | Asp | Glu | Glu | Asp | Ser | Gln |
| | 130 | | | | | 135 | | | | | 140 | | | | |
| Pro | Leu | Lys | Cys | Ile | Tyr | Gly | Arg | Pro | Gly | Gly | Pro | Val | Phe | Thr | Arg |
| 145 | | | | | 150 | | | | | 155 | | | | 160 | |
| Ser | Ala | Phe | Ala | Ala | Val | Leu | His | His | His | Gln | Asn | Pro | Glu | Phe | Tyr |
| | | | 165 | | | | | 170 | | | | | | 175 | |
| Asp | Glu | Ile | Lys | Ile | Glu | Leu | Pro | Thr | Gln | Leu | His | Glu | Lys | His | His |
| | | | 180 | | | | | 185 | | | | | 190 | | |
| Leu | Leu | Leu | Thr | Phe | Phe | His | Val | Ser | Cys | Asp | Asn | Ser | Ser | Lys | Gly |
| | | 195 | | | | | 200 | | | | | 205 | | | |
| Ser | Thr | Lys | Lys | Arg | Asp | Val | Val | Glu | Thr | Gln | Val | Gly | Tyr | Ser | Trp |
| | 210 | | | | | 215 | | | | | 220 | | | | |
| Leu | Pro | Leu | Leu | Lys | Asp | Gly | Arg | Val | Val | Thr | Ser | Glu | Gln | His | Ile |
| 225 | | | | | 230 | | | | | 235 | | | | 240 | |
| Pro | Val | Ser | Ala | Asn | Leu | Pro | Ser | Gly | Tyr | Leu | Gly | Tyr | Gln | Glu | Leu |
| | | | 245 | | | | | 250 | | | | | 255 | | |
| Gly | Met | Gly | Arg | His | Tyr | Gly | Pro | Glu | Ile | Lys | Trp | Val | Asp | Gly | Gly |

[illegible]

| | | |
|---|------|------|
| 690 | 695 | 700 |
| Ser Thr Pro Asn Ile Asn Ser Val Arg Asn Ala Asp Ser Arg Gly Ser | | |
| 705 | 710 | 715 |
| Leu Ile Ser Thr Asp Ser Gly Asn Ser Leu Pro Glu Arg Asn Ser Glu | | 720 |
| | 725 | 730 |
| Lys Ser Asn Ser Leu Asp Lys His Gln Gln Ser Ser Thr Leu Gly Asn | | 735 |
| | 740 | 745 |
| Ser Val Val Arg Cys Asp Lys Leu Asp Gln Ser Glu Ile Lys Ser Leu | | 750 |
| | 755 | 760 |
| Leu Met Cys Phe Leu Tyr Ile Leu Lys Ser Met Ser Asp Asp Ala Leu | | 765 |
| | 770 | 775 |
| Phe Thr Tyr Trp Asn Lys Ala Ser Thr Ser Glu Leu Met Asp Phe Phe | | 780 |
| 785 | 790 | 795 |
| Thr Ile Ser Glu Val Cys Leu His Gln Phe Gln Tyr Met Gly Lys Arg | | 800 |
| | 805 | 810 |
| Tyr Ile Ala Arg Thr Gly Met Met His Ala Arg Leu Gln Gln Leu Gly | | 815 |
| | 820 | 825 |
| Ser Leu Asp Asn Ser Leu Thr Phe Asn His Ser Tyr Gly His Ser Asp | | 830 |
| | 835 | 840 |
| Ala Asp Val Leu His Gln Ser Leu Leu Glu Ala Asn Ile Ala Thr Glu | | 845 |
| 850 | 855 | 860 |
| Val Cys Leu Thr Ala Leu Asp Thr Leu Ser Leu Phe Thr Leu Ala Phe | | 865 |
| | 870 | 875 |
| Lys Asn Gln Leu Leu Ala Asp His Gly His Asn Pro Leu Met Lys Lys | | 880 |
| | 885 | 890 |
| Val Phe Asp Val Tyr Leu Cys Phe Leu Gln Lys His Gln Ser Glu Thr | | 895 |
| | 900 | 905 |
| Ala Leu Lys Asn Val Phe Thr Ala Leu Arg Ser Leu Ile Tyr Lys Phe | | 910 |
| | 915 | 920 |
| Pro Ser Thr Phe Tyr Glu Gly Arg Ala Asp Met Cys Ala Ala Leu Cys | | 925 |
| 930 | 935 | 940 |
| Tyr Glu Ile Leu Lys Cys Cys Asn Ser Lys Leu Ser Ser Ile Arg Thr | | 945 |
| | 950 | 955 |
| Glu Ala Ser Gln Leu Leu Tyr Phe Leu Met Arg Asn Asn Phe Asp Tyr | | 960 |
| | 965 | 970 |
| Thr Gly Lys Lys Ser Phe Val Arg Thr His Leu Gln Val Ile Ile Ser | | 975 |
| | 980 | 985 |
| Val Ser Gln Leu Ile Ala Asp Val Val Gly Ile Gly Gly Thr Arg Phe | | 990 |
| | 995 | 1000 |
| Gln Gln Ser Leu Ser Ile Ile Asn Asn Cys Ala Asn Ser Asp Arg Leu | | 1005 |
| | 1010 | 1015 |
| Ile Lys His Thr Ser Phe Ser Ser Asp Val Lys Asp Leu Thr Lys Arg | | 1020 |
| 1025 | 1030 | 1035 |
| Ile Arg Thr Val Leu Met Ala Thr Ala Gln Met Lys Glu His Glu Asn | | 1040 |
| | 1045 | 1050 |
| Asp Pro Glu Met Leu Val Asp Leu Gln Tyr Ser Leu Ala Lys Ser Tyr | | 1055 |
| | 1060 | 1065 |
| Ala Ser Thr Pro Glu Leu Arg Lys Thr Trp Leu Asp Ser Met Ala Arg | | 1070 |
| | 1075 | 1080 |
| Ile His Val Lys Asn Gly Asp Leu Ser Glu Ala Ala Met Cys Tyr Val | | 1085 |
| | 1090 | 1095 |
| His Val Thr Ala Leu Val Ala Glu Tyr Leu Thr Arg Lys Glu Ala Val | | 1100 |
| 1105 | 1110 | 1115 |
| Gln Trp Glu Pro Pro Leu Leu Pro His Ser His Ser Ala Cys Leu Arg | | 1120 |

1125 1130 1135
 Arg Ser Arg Gly Gly Val Phe Arg Gln Gly Cys Thr Ala Phe Arg Val
 1140 1145 1150
 Ile Thr Pro Asn Ile Asp Glu Glu Ala Ser Met Met Glu Asp Val Gly
 1155 1160 1165
 Met Gln Asp Val His Phe Asn Glu Asp Val Leu Met Glu Leu Leu Glu
 1170 1175 1180
 Gln Cys Ala Asp Gly Leu Trp Lys Ala Glu Arg Tyr Glu Leu Ile Ala
 1185 1190 1195 1200
 Asp Ile Tyr Lys Leu Ile Ile Pro Ile Tyr Glu Lys Arg Arg Asp Phe
 1205 1210 1215
 Glu Arg Leu Ala His Leu Tyr Asp Thr Leu His Arg Ala Tyr Ser Lys
 1220 1225 1230
 Val Thr Glu Val Met His Ser Gly Arg Arg Leu Leu Gly Thr Tyr Phe
 1235 1240 1245
 Arg Val Ala Phe Phe Gly Gln Ala Ala Gln Tyr Gln Phe Thr Asp Ser
 1250 1255 1260
 Glu Thr Asp Val Glu Gly Phe Phe Glu Asp Glu Asp Gly Lys Glu Tyr
 1265 1270 1275 1280
 Ile Tyr Lys Glu Pro Lys Leu Thr Pro Leu Ser Glu Ile Ser Gln Arg
 1285 1290 1295
 Leu Leu Lys Leu Tyr Ser Asp Lys Phe Gly Ser Glu Asn Val Lys Met
 1300 1305 1310
 Ile Gln Asp Ser Gly Lys Val Asn Pro Lys Asp Leu Asp Ser Lys Tyr
 1315 1320 1325
 Ala Tyr Ile Gln Val Thr His Val Ile Pro Phe Phe Asp Glu Lys Glu
 1330 1335 1340
 Leu Gln Glu Arg Lys Thr Glu Phe Glu Arg Ser His Asn Ile Arg Arg
 1345 1350 1355 1360
 Phe Met Phe Glu Met Pro Phe Thr Gln Thr Gly Lys Arg Gln Gly Gly
 1365 1370 1375
 Val Glu Glu Gln Cys Lys Arg Arg Thr Ile Leu Thr Ala Ile His Cys
 1380 1385 1390
 Phe Pro Tyr Val Lys Lys Arg Ile Pro Val Met Tyr Gln His His Thr
 1395 1400 1405
 Asp Leu Asn Pro Ile Glu Val Ala Ile Asp Glu Met Ser Lys Lys Val
 1410 1415 1420
 Ala Glu Leu Arg Gln Leu Cys Ser Ser Ala Glu Val Asp Met Ile Lys
 1425 1430 1435 1440
 Leu Gln Leu Lys Leu Gln Gly Ser Val Ser Val Gln Val Asn Ala Gly
 1445 1450 1455
 Pro Leu Ala Tyr Ala Arg Ala Phe Leu Asp Asp Thr Asn Thr Lys Arg
 1460 1465 1470
 Tyr Pro Asp Asn Lys Val Lys Leu Leu Lys Glu Val Phe Arg Gln Phe
 1475 1480 1485
 Val Glu Ala Cys Gly Gln Ala Leu Ala Val Asn Glu Arg Leu Ile Lys
 1490 1495 1500
 Glu Asp Gln Leu Glu Tyr Gln Glu Glu Met Lys Ala Asn Tyr Arg Glu
 1505 1510 1515 1520
 Met Ala Lys Glu Leu Ser Glu Ile Met His Glu Gln Ile Cys Pro Leu
 1525 1530 1535
 Glu Glu Lys Thr Ser Val Leu Pro Asn Ser Leu His Ile Phe Asn Ala
 1540 1545 1550
 Ile Ser Gly Thr Pro Thr Ser Thr Met Val His Gly Met Thr Ser Ser

1555
Ser Ser Val Val
1570

1560

1565

<210> 159
<211> 540
<212> DNA
<213> Homo sapiens

<400> 159
gccggctctg ccattgtgctt actctgagcc acctaaccctc ggcgtgcttc agtttactca
60
tccgctcatc tgcagaatgg gtgatgtgtt cgggtacttcg tggcatacag gaaagtgcac
120
agcatgggtca gcctcagtga gaggtggcca gtggggagtg gtggccactg tacacctggc
180
acagcccaga gatgcatgtg ccactctgtt gtgtgcttca accaaggggc gctctggcag
240
ggcttgggtg ggacttccca aagggcattg aaaagttccc agtcaatgag atccatggag
300
acccatggga gtgggggtca gcccagcct aagaggaccc ccagccctgc cctgtgcccc
360
aggacacacc aggactgtc ccttgtcgcc ttcccagaca acctgtaccc tccaggccac
420
cagttctcgt ccattgacaaa gaaaggagcc ttctaaataa gtgcccgcga gaggtgcac
480
gttccctgc ccctccggg tggacctggg ttcaaagag aagctgccag tgcaacgcgt
540

<210> 160
<211> 110
<212> PRT
<213> Homo sapiens

<400> 160
Met Val Ser Leu Ser Glu Arg Trp Pro Val Gly Ser Gly Gly His Cys
1 5 10 15
Thr Pro Gly Thr Ala Gln Arg Cys Met Cys His Ser Val Val Cys Phe
20 25 30
Asn Gln Gly Ala Leu Trp Gln Gly Leu Gly Gly Thr Ser Gln Arg Ala
35 40 45
Trp Lys Ser Ser Gln Ser Met Arg Ser Met Glu Thr His Gly Ser Gly
50 55 60
Gly Gln Pro Gln Pro Lys Arg Thr Pro Ser Pro Ala Leu Cys Pro Arg
65 70 75 80
Thr His Gln Ala Leu Ser Leu Val Ala Phe Pro Asp Asn Leu Tyr Pro
85 90 95
Pro Gly His Gln Phe Ser Ser Met Thr Lys Lys Gly Ala Phe
100 105 110

<210> 161
<211> 351
<212> DNA
<213> Homo sapiens

<400> 161
 nnacgcgtac gtcttttcggc cgaagaagga acgtgggcag gggcctcctt cgctggccgc
 60
 cgcgcttggc tcgcagcgac gatgaagggc gacgacagca gcaagatcac ccacaagatc
 120
 gcccggggcga agcgcgaggg ccgcgtatgg tggagctttg agtacttccc gccgcgcacg
 180
 ccgcaggggca tgcagaattt gtatgaccgt atcgagcgca tgagtcagct gggccccgag
 240
 tttgtggaca ttacgtggaa tgccgggggc cggacgtcgg atatgacgac gcagctggtc
 300
 aagacggtgc atgcgtactt tgggtgctgag acgtgcatgc atctgacgtg c
 351

<210> 162
 <211> 117
 <212> PRT
 <213> Homo sapiens

<400> 162
 Xaa Arg Val Arg Leu Ser Ala Glu Glu Gly Thr Trp Ala Gly Ala Ser
 1 5 10 15
 Phe Ala Gly Arg Arg Ala Trp Leu Ala Ala Thr Met Lys Gly Asp Asp
 20 25 30
 Ser Ser Lys Ile Thr His Lys Ile Ala Arg Ala Lys Arg Glu Gly Arg
 35 40 45
 Val Trp Trp Ser Phe Glu Tyr Phe Pro Pro Arg Thr Pro Gln Gly Met
 50 55 60
 Gln Asn Leu Tyr Asp Arg Ile Glu Arg Met Ser Gln Leu Gly Pro Glu
 65 70 75 80
 Phe Val Asp Ile Thr Trp Asn Ala Gly Gly Arg Thr Ser Asp Met Thr
 85 90 95
 Thr Gln Leu Val Lys Thr Val His Ala Tyr Phe Gly Val Glu Thr Cys
 100 105 110
 Met His Leu Thr Cys
 115

<210> 163
 <211> 360
 <212> DNA
 <213> Homo sapiens

<400> 163
 gcgtgctcca tcggcacctt gcagatgggc gaattcgctg aaaacgtcgc cgggtggcgtc
 60
 gacacctaca ccttgcgtca gcccatcggc gtatgctcag gcatcactcc gttcaacttc
 120
 ccggcgatga ttccactgtg gatgttcccg atggcgattg cctgcggtaa cactttcgtg
 180
 ctcaaaccgt ccgaacaaga ccctctgtcg acgatgctgc tggtagaact ggcgctggaa
 240
 gccggtgtgc cggccggcgt gctcaacgtg gtgcacggcg gcaaggatgt ggtggatgctg
 300

ctgtgcaccc ataaagatat caaggcagtt tctttcgtcg gttcgaccgc cgttgggtacc
360

<210> 164

<211> 120

<212> PRT

<213> Homo sapiens

<400> 164

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Ala | Cys | Ser | Ile | Gly | Thr | Leu | Gln | Met | Gly | Glu | Phe | Ala | Glu | Asn | Val |
| 1 | | | | 5 | | | | | 10 | | | | | 15 | |
| Ala | Gly | Gly | Val | Asp | Thr | Tyr | Thr | Leu | Arg | Gln | Pro | Ile | Gly | Val | Cys |
| | | | 20 | | | | | 25 | | | | | 30 | | |
| Ala | Gly | Ile | Thr | Pro | Phe | Asn | Phe | Pro | Ala | Met | Ile | Pro | Leu | Trp | Met |
| | | 35 | | | | | 40 | | | | | 45 | | | |
| Phe | Pro | Met | Ala | Ile | Ala | Cys | Gly | Asn | Thr | Phe | Val | Leu | Lys | Pro | Ser |
| | 50 | | | | | 55 | | | | | 60 | | | | |
| Glu | Gln | Asp | Pro | Leu | Ser | Thr | Met | Leu | Leu | Val | Glu | Leu | Ala | Leu | Glu |
| 65 | | | | 70 | | | | | | 75 | | | | 80 | |
| Ala | Gly | Val | Pro | Ala | Gly | Val | Leu | Asn | Val | Val | His | Gly | Gly | Lys | Asp |
| | | | 85 | | | | | 90 | | | | | | 95 | |
| Val | Val | Asp | Ala | Leu | Cys | Thr | His | Lys | Asp | Ile | Lys | Ala | Val | Ser | Phe |
| | | 100 | | | | | | 105 | | | | | 110 | | |
| Val | Gly | Ser | Thr | Ala | Val | Gly | Thr | | | | | | | | |
| | | 115 | | | | | 120 | | | | | | | | |

<210> 165

<211> 728

<212> DNA

<213> Homo sapiens

<400> 165

gctagcagcc ttcacccctcc tagaggggca ggctcggcga caaggggagg ggggtgccccg
60
tcccagcgag ggacgccccg ggctgggggt gccgggtcag cccggggcaa cagcttcacc
120
aagtttgga accgcaacgt cttcatgaag gacaacagct cttcttccag cacagactcc
180
cgctcccgt cctcctccag gtccccgacg cgccacttcc gcagaagtga ctcccactca
240
gactccgaca gctcctactc agggaatgag tgtcaccctg tgggcccag gaaccgccc
300
cctaagggcc ggggagggtc agggggccat atggatcggg gccgaggcag ggcgcagcgt
360
gggaagaggc acgatctggc gccaccaaag cgcagtcgaa agaagatggc ggcgctggag
420
tgtgaggacc cggagcgaga gctgaagaag cagaagcggg cagcccgtt ccagcacgga
480
cactccccgc gcctgcgcct cgagcccctg gtgctgcaga tgagcagcct ggagagcagt
540
ggggctgacc ctgactggca ggagctgcag atcgtgggca cctgccctga catcaccaag
600
cactacctgc gcctcacctg tgccccgac cgtccaccg tgcgccctgt ggcattccct
660

gtggcaggtt ttgaaaaagt cgctgtgcat ggtcaagtgc cactggaaaag agaagcagga
 720
 ctacgcgt
 728

<210> 166
 <211> 242
 <212> PRT
 <213> Homo sapiens

<400> 166
 Ala Ser Ser Leu His Pro Pro Arg Gly Ala Gly Ser Ala Thr Arg Gly
 1 5 10 15
 Gly Gly Ala Pro Ser Gln Arg Gly Thr Pro Gly Ala Gly Gly Ala Gly
 20 25 30
 Arg Ala Arg Gly Asn Ser Phe Thr Lys Phe Gly Asn Arg Asn Val Phe
 35 40 45
 Met Lys Asp Asn Ser Ser Ser Ser Thr Asp Ser Arg Ser Arg Ser
 50 55 60
 Ser Ser Arg Ser Pro Thr Arg His Phe Arg Arg Ser Asp Ser His Ser
 65 70 75 80
 Asp Ser Asp Ser Ser Tyr Ser Gly Asn Glu Cys His Pro Val Gly Arg
 85 90 95
 Arg Asn Pro Pro Pro Lys Gly Arg Gly Gly Arg Gly Ala His Met Asp
 100 105 110
 Arg Gly Arg Gly Arg Ala Gln Arg Gly Lys Arg His Asp Leu Ala Pro
 115 120 125
 Thr Lys Arg Ser Arg Lys Lys Met Ala Ala Leu Glu Cys Glu Asp Pro
 130 135 140
 Glu Arg Glu Leu Lys Lys Gln Lys Arg Ala Ala Arg Phe Gln His Gly
 145 150 155 160
 His Ser Arg Arg Leu Arg Leu Glu Pro Leu Val Leu Gln Met Ser Ser
 165 170 175
 Leu Glu Ser Ser Gly Ala Asp Pro Asp Trp Gln Glu Leu Gln Ile Val
 180 185 190
 Gly Thr Cys Pro Asp Ile Thr Lys His Tyr Leu Arg Leu Thr Cys Ala
 195 200 205
 Pro Asp Pro Ser Thr Val Arg Pro Val Ala Phe Pro Val Ala Gly Phe
 210 215 220
 Glu Lys Val Ala Val His Gly Gln Val Pro Leu Glu Arg Glu Ala Gly
 225 230 235 240
 Leu Arg

<210> 167
 <211> 510
 <212> DNA
 <213> Homo sapiens

<400> 167
 nnacgcgtgg aaccagaact caggcccgtg tgaggagtct ggtttggaac acacggggcc
 60
 gcaacacaga attgtcaggt cctgtgccgt gaccaccaac cctcggggcca tgccaggtgc
 120

tgggtgagggg caggtgggtc ccgccaggcg cctgctggcc tgaccgcact ccgtccacag
 180
 gtccctcatgg ggcctcctcg gctgggcttc gtgtccgcct acctctcaca gccactgctc
 240
 gatggctttg ccatgggggc ctccgtgacc atcctgacct cgcagctcaa acacctgctg
 300
 ggcggtcgga tcccgcggca ccagggggcc ggcatggtgg tcctcacatg gctgagcctg
 360
 ctgcgcggcg ccgggcaggc caacgtgtgc gacgtgggtc ccagcacggg gtgcctggcg
 420
 gtgctgctag ccgcgaagga gctctcagac cgctaccgac accgcctgag ggtgccgctg
 480
 cccacggagc tgctgggtcat cgtgggtggc
 510

<210> 168

<211> 128

<212> PRT

<213> Homo sapiens

<400> 168

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Gly | Ala | Gly | Gly | Ser | Arg | Gln | Ala | Pro | Ala | Gly | Leu | Thr | Ala | Leu | Arg |
| 1 | | | | 5 | | | | | 10 | | | | | 15 | |
| Pro | Gln | Val | Leu | Met | Gly | Val | Leu | Arg | Leu | Gly | Phe | Val | Ser | Ala | Tyr |
| | | 20 | | | | | | 25 | | | | | 30 | | |
| Leu | Ser | Gln | Pro | Leu | Leu | Asp | Gly | Phe | Ala | Met | Gly | Ala | Ser | Val | Thr |
| | | 35 | | | | | 40 | | | | | 45 | | | |
| Ile | Leu | Thr | Ser | Gln | Leu | Lys | His | Leu | Leu | Gly | Val | Arg | Ile | Pro | Arg |
| | 50 | | | | | 55 | | | | | 60 | | | | |
| His | Gln | Gly | Pro | Gly | Met | Val | Val | Leu | Thr | Trp | Leu | Ser | Leu | Leu | Arg |
| 65 | | | | | 70 | | | | | 75 | | | | 80 | |
| Gly | Ala | Gly | Gln | Ala | Asn | Val | Cys | Asp | Val | Val | Thr | Ser | Thr | Val | Cys |
| | | | 85 | | | | | 90 | | | | | | 95 | |
| Leu | Ala | Val | Leu | Leu | Ala | Ala | Lys | Glu | Leu | Ser | Asp | Arg | Tyr | Arg | His |
| | | 100 | | | | | 105 | | | | | | 110 | | |
| Arg | Leu | Arg | Val | Pro | Leu | Pro | Thr | Glu | Leu | Leu | Val | Ile | Val | Val | Ala |
| | | 115 | | | | | 120 | | | | | | 125 | | |

<210> 169

<211> 537

<212> DNA

<213> Homo sapiens

<400> 169

gaattccacc gcatgtcgtg tctggacgta tgtaggtcgc ggtagtgtgc gaccgccggg
 60
 gccttaaagg agagcgggca tcggcgttgc agtacgagag gggaaggtgt gcggatactt
 120
 attgtcgggtg cggcatcgtc catccacacc gttcgatggg tcaatggact ggtcaagcgg
 180
 ggtcacgagg ttcacctggc atcagtccat ccggcggggc gtcactccat tgatccccga
 240
 gttcggatcc acctggcccc acacggcggg aaggcaaaat acgtcgtcaa tgccggctgg
 300

ctgcgatcag tggcggctgg ggtgcaacct gacatcgta acgtccacta tgcgaccggt
 360
 tatggctctgc tcgctcgtct tgcccatatt gacgccccga cgctgctgtc ggtgtgggga
 420
 agtgacgttt acgattcccc ccgggcaaatt cccctcatgc gtcacatggt ccgatccaac
 480
 ttggtctcag ctactcggat cgcacgcaca agccactgca tggcgcgtgt cacgcgt
 537

<210> 170

<211> 164

<212> PRT

<213> Homo sapiens

<400> 170

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Cys | Ala | Thr | Ala | Gly | Ala | Leu | Lys | Glu | Ser | Gly | His | Arg | Arg | Cys | Ser |
| 1 | | | | 5 | | | | | 10 | | | | | 15 | |
| Thr | Arg | Gly | Glu | Gly | Val | Arg | Ile | Leu | Ile | Val | Gly | Ala | Ala | Ser | Ser |
| | | 20 | | | | | | 25 | | | | | 30 | | |
| Ile | His | Thr | Val | Arg | Trp | Val | Asn | Gly | Leu | Val | Lys | Arg | Gly | His | Glu |
| | | 35 | | | | | 40 | | | | | 45 | | | |
| Val | His | Leu | Ala | Ser | Val | His | Pro | Ala | Gly | Arg | His | Ser | Ile | Asp | Pro |
| | 50 | | | | | 55 | | | | 60 | | | | | |
| Arg | Val | Arg | Ile | His | Leu | Ala | Pro | His | Gly | Gly | Lys | Ala | Lys | Tyr | Val |
| 65 | | | | 70 | | | | | 75 | | | | | 80 | |
| Val | Asn | Ala | Gly | Trp | Leu | Arg | Ser | Val | Ala | Ala | Gly | Val | Gln | Pro | Asp |
| | | 85 | | | | | | 90 | | | | | 95 | | |
| Ile | Val | Asn | Val | His | Tyr | Ala | Thr | Gly | Tyr | Gly | Leu | Leu | Ala | Arg | Leu |
| | | 100 | | | | | | 105 | | | | | 110 | | |
| Ala | His | Ile | Asp | Ala | Pro | Thr | Leu | Leu | Ser | Val | Trp | Gly | Ser | Asp | Val |
| | | 115 | | | | | 120 | | | | | 125 | | | |
| Tyr | Asp | Ser | Pro | Arg | Ala | Asn | Pro | Leu | Met | Arg | His | Met | Val | Arg | Ser |
| | 130 | | | | | 135 | | | | | 140 | | | | |
| Asn | Leu | Val | Ser | Ala | Thr | Arg | Ile | Ala | Ser | Thr | Ser | His | Cys | Met | Ala |
| 145 | | | | | 150 | | | | | 155 | | | | 160 | |
| Arg | Val | Thr | Arg | | | | | | | | | | | | |

<210> 171

<211> 391

<212> DNA

<213> Homo sapiens

<400> 171

ctagacaagc tcgcgcgggt gggcttcgac actcttggtc tacagacctt cctaactgcg
 60
 ggggagaagg agtcccgcgc atggacgatt cacaagggcg acaccgcccc tgaggctgct
 120
 ggcgtcatcc ataccgactt ccagaagggg ttcacaaagg cccaggtggt gtccttcggc
 180
 gaccttggtg aatttgccgg cgaaaaggag gccagggctg ctgggaagct gcggttgga
 240
 ggcaaggagt acgttatgca ggacgggtgac gtagtggaat tccgatttaa cgtgtagctc
 300

tggtttgata cttacttggc ttaaccgcat ctgagatccg tcatatcttt ggcgtagcct
 360
 tattggtatg aataacatgc cgtagccaaa g
 391

<210> 172
 <211> 98
 <212> PRT
 <213> Homo sapiens

<400> 172
 Leu Asp Lys Leu Ala Arg Val Gly Phe Asp Thr Leu Gly Leu Gln Thr
 1 5 10 15
 Phe Leu Thr Ala Gly Glu Lys Glu Ser Arg Ala Trp Thr Ile His Lys
 20 25 30
 Gly Asp Thr Ala Pro Glu Ala Ala Gly Val Ile His Thr Asp Phe Gln
 35 40 45
 Lys Gly Phe Ile Lys Ala Gln Val Val Ser Phe Gly Asp Leu Val Glu
 50 55 60
 Phe Gly Gly Glu Lys Glu Ala Gln Ala Ala Gly Lys Leu Arg Leu Glu
 65 70 75 80
 Gly Lys Glu Tyr Val Met Gln Asp Gly Asp Val Val Glu Phe Arg Phe
 85 90 95
 Asn Val

<210> 173
 <211> 309
 <212> DNA
 <213> Homo sapiens

<400> 173
 ccatggagtg tcccttgtgc gagcattttg agagctatac caacacccat ccctgcagg
 60
 cccagagccg agccatttct caggagagca ggaagggagc aggccgaggg gtgctcccag
 120
 ccagccccgg aaccgaggt ctggggacgc agccgaccag cctccttgt ctgggcctct
 180
 gtttcctctt cgacacaggg aagcagggag gggccgatca gcgacttagg cctgttggt
 240
 gtggtgggggt cccctgcgtt tctgggaagc cacggaccct gggatgtacc tgggtttcat
 300
 tcgcagtga
 309

<210> 174
 <211> 102
 <212> PRT
 <213> Homo sapiens

<400> 174
 Met Glu Cys Pro Leu Cys Glu His Phe Glu Ser Tyr Thr Asn Thr His
 .1 5 10 15
 Pro Cys Arg Ser Gln Ser Arg Ala Ile Ser Gln Glu Ser Arg Lys Gly

| | | | | | |
|-------------|-------------|------------|-------------|------------|-------------|
| 400> 175 | | | | | |
| nnacttttttt | tttttttttt | catttatgct | atggagaaaac | cagcatggag | atgtcatggg |
| 60 | | | | | |
| agagcatgca | caggccccgc | cctagggagt | ggtgatgtgt | ttggggaggt | gcttgtttcc |
| 120 | | | | | |
| aggttccatcc | cacacgttgt | ccagttggat | cctatggcag | gctggctgtg | gctttctctc |
| 180 | | | | | |
| tcctgcttct | cttcctcctc | cagataaggg | tctgcaggat | cttctgctta | gcaagtgggtg |
| 240 | | | | | |
| gccaaaggact | ggtggatggg | tggttggaag | cagcgcacat | gctccacagt | ggaactgtct |
| 300 | | | | | |
| gtctccacgg | acttcatgta | tttgttcagg | atggcaaaaa | cctcattggt | caagatctga |
| 360 | | | | | |
| tacttctga | tccggtcggc | catcttcttc | aggggcacat | tcttaatgat | ttcatccttc |
| 420 | | | | | |
| ccgtctgcc | tctgcacttt | tagcaggtgg | taacagaagt | cgaacaggtc | aaagcgacgc |
| 480 | | | | | |
| tgctggccca | gcaggacaat | gatggagcaa | ccagcccagt | tcaagccatc | gccgaaacac |
| 540 | | | | | |
| tgctcagctg | tgaactcggt | ggttcccaca | gggatgcagt | acacgaactg | catggcgctc |
| 600 | | | | | |
| cacagccggg | ggaactccac | acactcatcg | acgtgcatga | cgccattggg | gggcgggtggg |
| 660 | | | | | |
| ccccgccaga | tggggctctg | caggtagctc | cgaatgcggg | tcaggatgac | ctcaaacatg |
| 720 | | | | | |
| gacaggccac | agcacagccg | ctccttggtc | aggagggtcac | cctcgcgagc | aatggcgatt |
| 780 | | | | | |
| tgctgagggg | tccccagccg | ctcgatcaga | gggaccaggt | ggagcggggc | atacttggct |
| 840 | | | | | |
| tccagacggt | tcatttttggc | atcaagtctc | tccccctctt | tcacatggac | tcgcggcaag |
| 900 | | | | | |
| atgttcttga | aaggagccgc | gtgcagcagg | tcacacactt | cttctaaaga | cagagctctg |
| 960 | | | | | |
| ctcgatgagg | aggcagaaga | ggatggcatt | gccacttcc | ctcaggctct | ggaagcacgt |
| 1020 | | | | | |
| ctgttttgag | ctctgcgtac | tcaatgatgt | ccttcagctg | gtggtggaag | aactccagga |
| 1080 | | | | | |

tccttgggga gccataactca tgtcggggca agcggcatat cttgggcatc acctctatca
1140
gtgtttttcac atactggaga atggttcctt ggagcaagct cttcacaatc ttttagcagt
1200
cctccatgac cacagcgatg ccctgataac ccaggagtct gcagatagtc ttgaaagtgt
1260
ggagggtocca cgaagtccg gtagctgccg taaatgctgg agtaggcca gttcaaagcc
1320
ttggatccat gcagatactg aggtctgtgca ttaggtctgt tatctctttg aaattcctga
1380
gaaaatggta aactgtccg acaaaaccgg ttggtagagc cgtttagca gtagttgggc
1440
aggaagtc atgttagctc ccagaagacg tgcagggtga tcctcccgta gggcgctgac
1500
acgttgtggg tggcctcccg gaacatggcg tcgaagccgt ccagcgtcag gtaccggctc
1560
agcagcttgt gggcatgagc gttgatttcc aacaggccat ccagctcaac tatggaggtc
1620
aaatcttcac tttcaaatcg tccaatcgcc agttctaggg acttatacat ggctgctgag
1680
acgcgctggg tgatcagacg attgaggtct attgatctgc cgaggagctg cacatgcctc
1740
tgcttcagca gcgtctcgta gcggttagac ggcgggaggt ggatcgtggc tccttgatc
1800
ttgcattctg atcgtaaccg tttatcaaga agcaaacttc ctgccataac cttataatag
1860
gcaaatatct ggtctgccag cttgtagaca aactgatcaa aacacagggt cacctcagct
1920
tctatctcat cgtacaggaa ctgcttttta aacttgggtca gagcatagta ggcgctgtca
1980
ttgtacagat ccaggagta gagcacgtac tccatcatgg aaggttcttt ggtttccagg
2040
atatgggtccg ttagaatcca gggcatcgac atctcaatgg ggaactggat cctcctgccc
2100
atggtcagct ccaggaagaa ctctcggaac cacagctgcg aaaggtcaca gactgctgc
2160
agcgtttcac tgaaatttat caagtgagt tagaagaatg actctcgatg aaagtccctc
2220
atggcgagga caatgggtcc atccaggctg ctctcaggg tcttcttggg gccgcttttg
2280
tctgcaatga gtgattcaag catggttctc accatgtaca gctgtgtgct ggatggcccc
2340
acagcacgc ggggcacctt gatataaat ccacctttgg ggtccttctc ccctctcaag
2400
catgggtcat tagggggctc tcgccctccc tcccagtcac agatgggtct tcgaattgcc
2460
tgtaggagcg tgatgaggac attcttcttc ttcgtaaccg cctgccgcag gggctcacgc
2520
agcgtcacct gggcgaagtc ctgcaatgcc gcgtagatgg tgttcctgat ggctgggtg
2580
aagacgctct ccatcctgcc catgagcacc tgcaggcctt tgatcatggc gatcacctca
2640
acgaaggcaa atttttcctc actggtgtaa ttgtagcgtg tggctctctc atattcctcc
2700

gcgggtgccag gacagtcctt gttgcagaac ttgtctgtgg gatgaaccag cttccaagag
2760
tacacctcca tgacgtgggc gctccacttg gatagaagct gcagaccccg cagggctagg
2820
tcgaagagct cgcgatactc ctcgtctgac ttctggctgt ccagccctga gcccgtcacc
2880
acctcactgt tgctgtagcg agcgagctcg gagatgaagc ggatgtgggc atcccggatc
2940
tgaaccatct gctcgagat attgtactgg gggctgatgc tgctctgggt gcacgtccat
3000
cgagatttat tttcctcgta gtgggcgctg gtcttgatat atcttgccag ttctatttgc
3060
atgtccccaa atagtggaaac cacctggagt tgcttgaagt acttgctgat tttggataag
3120
tttattcttt tcttggcatc caacttatag atgttactga cactcccatc catcaggtag
3180
agaccaaata ccatgacttt gagaagcatg tgtttctcac tgggcgtcaa atacatcctg
3240
ttctcgtagt aatccacaca cagattcaca atatctgcca ggagctcttc gtagccagaa
3300
atcattcga gctgctgctg cagagactgt gtgatcttgt tatgattggc caggaacatg
3360
gacagattct gcgattcctg gatggactgt ggatctgcca ttttacgtaa aaactgtgct
3420
gccctcttgt aggcagagtg gtcgttcttg acgctgcaat tcatgttctt cagatcgcc
3480
agcacagcga acatgttgat gaatttgccc agtgtgatca ggtaggcttc tgacacgaag
3540
tccttcctgc gctcggcatg gcacagccgc ttcacctcgc tgcagaaccg ctcgatggcc
3600
ttgcgctgaa aatacatgaa cttcatgagc ttggtgacct ccggctccag cacctctact
3660
gtcttctcat agatctctac tcggttgggc tgctcgttgc atttcacctg gggaatggcc
3720
cgggaacagc tgcgccaggt gtacagcatg accgcatact catgtccttc ctccagcatc
3780
tcattcatgc tggagtggac tgtagcctgc tcaatgtacc ttgcaatgcc cgtgacaaat
3840
gcattcctgt cctcaaagtt tgtgtcaaag ttagcctggt acatgatgga ggaaggtgga
3900
ggctcgatgc atggctgctg gtcggggagg ggaagctctt caagcaggtc cacgttggac
3960
agggcatctt ccagggtgac gtgcgtgggc atggctgcag tttctgtatt ctgcactggc
4020
tctgcggcgg gtccagggga ggggcctgag gctaccgcag ggccccgcgc cggggcgcag
4080
gatgctctgc cccgctccgc tgcgccgcca gtgcagaata cagaaactgc agccatgacc
4140
acgcacgtca ccctggaaga tgccctgtcc aacgtggacc tgcttgaaga gcttcccctc
4200
cccgaccagc agccatgcat cgagcctcca ccttctcca tcatgtacca ggctaacttt
4260
gacacaaact ttgaggacag gaatgcattt gtcacgggca ttgcaaggta cattgagcag
4320

gctacagtcc actccagcat gaatgagatg ctggaggaag gacatgagta tgcggtcatg
4380
ctgtacacct ggcgagctg ctccccggcc atccccacagg tgaaatgtaa cgagcagcct
4440
aacagagtgg aaatctacga gaaaaccgtg gaggttctg agcctgaggt cacaaaactg
4500
atgaatttca tgtacttcca gagaaatgcc attgagcggt tctgcgggga agtgaggcgc
4560
ctgtgccatg ccgagaggag gaaggacttc gtgtcagaag cctacctgat cactctgggc
4620
aaattcatca acatgttcgc tgtgtctggac gagctgaaga acatgaagtg cagtgtgaag
4680
aacgaccact cagcgtacaa gagggccgct cagtttttac gtaaaatggc agatccacag
4740
tccatccagg aatcgcagaa tctgtccatg ttctctggcca atcataacaa gatcacacag
4800
tctctgcagc agcagctcga agtgatttct ggctacgaag agctcctggc agatattgtg
4860
aatctgtgtg tggattacta cgagaacagg atgtatttga cggccagtga gaaacacatg
4920
cttctcaaag tcatgggatt tggctctgtac ctgatggatg ggagtgtcag taacatctat
4980
aagttggatg ccaagaaaag aataaactta tccaaaatcg acaagtactt caagcaactc
5040
caggtgggtc cgctatttgg ggacatgcaa atagaactgg caagatatat caagaccagc
5100
gcccactacg aggaaaataa atctcgatgg acgtgcacat cctccggcag cagccctcag
5160
tacaacatct gcgagcagat gatccagatc cgcgaggacc acatgcgctt catttcggag
5220
ctggcgcgct acagcaacag cgaggtggtc acgggctcgg gccgccagga ggcccagaag
5280
acggacgcgg agtaccgcaa gctcttcgac ctggcgctgc agggcctgca gctgttgtcg
5340
cagtggagcg cgcacgtgat ggaagtgtat tcctggaagc ttgtgcaccc caccgacaag
5400
tactccaaca aggactgccc cgacagcgct gaagagtacg agcgtgccac gcgctacaac
5460
tacaccagcg aggagaagtt tgccctagtg gaggtgatcg ccatgatcaa aggcctgcag
5520
gtgctgatgg gcaggatgga gagcgtgttc aaccacgcca tccggcacac cgtctatgcc
5580
gcactgcagg acttctccca ggtgaccctt agggagccgc tgcggcaggc catcaagaag
5640
aagaagaacg tcatccagag tgtcctgcag gccatcagga agaccgtgtg tgactgggag
5700
acggggcatg agcccttcaa tgaccagcc ttgcggggcg agaaggaccc caagagcggc
5760
ttcgacataa aagtaccagc ccgcgccgtg ggaccctcca gcactcagct ttacatggtg
5820
agaaccatgc tagagtccct cattgcagac aaaagtgggt ccaagaaaac cttgagaagt
5880
agccttgagg ggcccacat attggacata gaaaaatttc atcgagagtc attcttctac
5940

actcacttga taaatttcag tgaacgctg cagcagtgtg gtgacctttc gcagctgtgg
6000
ttccgagagt ttttcttga gctgaccatg ggcaggagga tccagttccc cattgagatg
6060
tcgatgccct ggatcctgac ggaccacatc ctggagacca aggaggcatc gatgatggag
6120
tacgtgtctt actccctgga cctgtacaat gacagcgccc actacgcgct caccaggttc
6180
aacaagcagt tctgttacga cgaaattgag gccgaggtga atctatgttt tgaccaattt
6240
gtttacaagc tagcagacca gatatttgcc tattataagg ttatggcagg aagtttgctt
6300
cttgataaac ggttacgatc agaattgcaag aatcaggagga ccacgatcca cctcccgccg
6360
tctaaccgct acgagacgct gctgaagcag aggcattgtg agctcctcgg cagatcaata
6420
gacctcaatc gtctgatcac ccagcgctc tcagcagcca tgtataagtc cctagaactg
6480
gcgattggac gatttgaaag tgaagatttg acctccatag ttgagctgga tggcctgttg
6540
gaaatcaacc gcatgaccca caagctgctg agccggtacc tgacgctgga cggcttcgac
6600
gccattgttc gggaggccaa ccacaacgtg tcagcgccct acgggaggat caccctgcac
6660
gtcttctggg agtcaacta tgacttctg cccaactact gctacaacgg ctctaccaac
6720
cggtttgttc ggacagtgtt accattttct caggaatttc aaagagataa gcagccta
6780
gcacagcctc agtatctgca tggatccaag gctttgaact tggcctactc cagcatttac
6840
ggcagctacc ggaacttcgt gggacctcca cactttcaag tcattctgcc gcttctcggc
6900
taccagggta tcgccgtggt catggaggag ctgctgaagg tcgtcaagag cctgctgcaa
6960
ggcacaatcc tgagtagct gaagacgctg atggagggtga tgcccaagat ctgccgctg
7020
ccccggcagc agtacggctc tctgtgtatc ctggagtctt tccaccacca gctgaaggac
7080
atcgtggagt acgcagagct gaagacggtg tgcttccaga acctgcggga ggtggggaac
7140
gccatcctct tctgcctgct catcgagcag agcctgtctt tagaagaagt gtgtgacctg
7200
ctgcacgagg ctcctttcca gaacatcttg ccgcgagtc atgtgaaaga gggggagaga
7260
cttgatgcca aaatgaaaag actagaatca aagtacgccc cgctgcatct tgtccactg
7320
attgaaagac tggggacccc tcagcaaatt gccatcgcaa gagaggggga cctgctgaca
7380
aaggagcgcc tctgctgagg cctgtccatg tttgaggtca tctgacacg gatccggagc
7440
tttctggatg acccatctg gcgaggccct ctgcccagca atggggtcac gcatgtggac
7500
gagtgtgtgg agtttcacag actgtggagt gccatgcagt ttgtctactg cattcccggtg
7560

gggacacacg agttcacagt cgagcagtcg tttggtgatg ggctacactg ggctggctgt
 7620
 atgatcatcg tacttcttgg gcagcagcgg cgttttgctg tgctggattt ctgctaccat
 7680
 ctacttaaag tccagaaaca tgatggcaaa gatgagatta ttaaaaatgt gcctttgaag
 7740
 aagatggtgg agagaattcg caagttccag attctcaatg atgagatcat caccatcctg
 7800
 gataagtacc tgaagtcagg cgacggggag ggcacgccag tggagcatgt gcgctgcttc
 7860
 cagccgcca tccaccagtc cctcgccagc agctgagggc acgcgctgca ctccgtaact
 7920
 caacatggca tgcctttctc tccgtaaact atttagtgag attttttaggg actatttttc
 7980
 agtatctctg tacctgttaa aggggggtgct ttctgatcta aaaacttaat ttataaaat
 8040
 tgacttattt ttctagacta aaattgtata tgcttttggt aattaggaac tcttgagaat
 8100
 attggtgctt gattgttgcc atcacgttcc taaaaaattg tttttctatg ggatgttctg
 8160
 gcagctgtgt cataaaatgc tgctgggttc attcattcat tccataagaa acttaatacc
 8220
 agcaaatgca ttaaatccct tgccagttac cattaactgt aactatttag cttttgttta
 8280
 gggatctttc tgatggtctt ttatgagcaa tcttagttct aagtcattgt tcccatcctt
 8340
 tttttgtgtg ttccagaaaa tagtgaactt gattcccctg cttccactaa atccagttgt
 8400
 gacaaaatct aacgtgacat cagatcgaaa ggttatagaa ataaaactaa tgagatctaa
 8460
 aaaaaaaaaa aaaaaaaaaa aaaa
 8484

<210> 176
 <211> 1393
 <212> PRT
 <213> Homo sapiens

<400> 176
 Met Ala Arg Glu Gln Leu Arg Gln Val Tyr Ser Met Thr Ala Tyr Ser
 1 5 10 15
 Cys Pro Ser Ser Ile Ser Phe Met Leu Glu Trp Thr Val Ala Cys
 20 25 30
 Ser Met Tyr Leu Ala Met Pro Val Thr Asn Ala Phe Leu Ser Ser Lys
 35 40 45
 Phe Val Ser Lys Leu Ala Trp Tyr Met Met Glu Glu Gly Gly Gly Ser
 50 55 60
 Met His Gly Cys Trp Ser Gly Arg Gly Ser Ser Ser Ser Arg Ser Thr
 65 70 75 80
 Leu Asp Arg Ala Ser Ser Arg Val Thr Cys Val Val Met Ala Ala Val
 85 90 95
 Ser Val Phe Cys Thr Gly Ser Ala Ala Gly Pro Gly Glu Gly Pro Glu
 100 105 110
 Ala Thr Ala Gly Pro Arg Ala Gly Ala Gln Asp Ala Leu Pro Arg Ser

| | | | | | | | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|--|--|--|--|--|--|--|
| | | 115 | | | | | | | 120 | | | | | | 125 | | | | | | | |
| Ala | Ala | Pro | Pro | Val | Gln | Asn | Thr | Glu | Thr | Ala | Ala | Met | Thr | Thr | His | | | | | | | |
| | 130 | | | | | 135 | | | | | 140 | | | | | | | | | | | |
| Val | Thr | Leu | Glu | Asp | Ala | Leu | Ser | Asn | Val | Asp | Leu | Leu | Glu | Glu | Leu | | | | | | | |
| 145 | | | | | 150 | | | | | 155 | | | | | 160 | | | | | | | |
| Pro | Leu | Pro | Asp | Gln | Gln | Pro | Cys | Ile | Glu | Pro | Pro | Pro | Ser | Ser | Ile | | | | | | | |
| | | | | 165 | | | | | 170 | | | | | | 175 | | | | | | | |
| Met | Tyr | Gln | Ala | Asn | Phe | Asp | Thr | Asn | Phe | Glu | Asp | Arg | Asn | Ala | Phe | | | | | | | |
| | | | 180 | | | | | 185 | | | | | 190 | | | | | | | | | |
| Val | Thr | Gly | Ile | Ala | Arg | Tyr | Ile | Glu | Gln | Ala | Thr | Val | His | Ser | Ser | | | | | | | |
| | | 195 | | | | | 200 | | | | | 205 | | | | | | | | | | |
| Met | Asn | Glu | Met | Leu | Glu | Glu | Gly | His | Glu | Tyr | Ala | Val | Met | Leu | Tyr | | | | | | | |
| | 210 | | | | | 215 | | | | | 220 | | | | | | | | | | | |
| Thr | Trp | Arg | Ser | Cys | Ser | Arg | Ala | Ile | Pro | Gln | Val | Lys | Cys | Asn | Glu | | | | | | | |
| 225 | | | | | 230 | | | | | 235 | | | | | 240 | | | | | | | |
| Gln | Pro | Asn | Arg | Val | Glu | Ile | Tyr | Glu | Lys | Thr | Val | Glu | Val | Leu | Glu | | | | | | | |
| | | | | 245 | | | | | 250 | | | | | | 255 | | | | | | | |
| Pro | Glu | Val | Thr | Lys | Leu | Met | Asn | Phe | Met | Tyr | Phe | Gln | Arg | Asn | Ala | | | | | | | |
| | | | 260 | | | | | 265 | | | | | 270 | | | | | | | | | |
| Ile | Glu | Arg | Phe | Cys | Gly | Glu | Val | Arg | Arg | Leu | Cys | His | Ala | Glu | Arg | | | | | | | |
| | | 275 | | | | | 280 | | | | | 285 | | | | | | | | | | |
| Arg | Lys | Asp | Phe | Val | Ser | Glu | Ala | Tyr | Leu | Ile | Thr | Leu | Gly | Lys | Phe | | | | | | | |
| | 290 | | | | | 295 | | | | | 300 | | | | | | | | | | | |
| Ile | Asn | Met | Phe | Ala | Val | Leu | Asp | Glu | Leu | Lys | Asn | Met | Lys | Cys | Ser | | | | | | | |
| 305 | | | | | 310 | | | | | 315 | | | | | 320 | | | | | | | |
| Val | Lys | Asn | Asp | His | Ser | Ala | Tyr | Lys | Arg | Ala | Ala | Gln | Phe | Leu | Arg | | | | | | | |
| | | | | 325 | | | | | 330 | | | | | | 335 | | | | | | | |
| Lys | Met | Ala | Asp | Pro | Gln | Ser | Ile | Gln | Glu | Ser | Gln | Asn | Leu | Ser | Met | | | | | | | |
| | | 340 | | | | | | 345 | | | | | 350 | | | | | | | | | |
| Phe | Leu | Ala | Asn | His | Asn | Lys | Ile | Thr | Gln | Ser | Leu | Gln | Gln | Gln | Leu | | | | | | | |
| | | 355 | | | | | 360 | | | | | 365 | | | | | | | | | | |
| Glu | Val | Ile | Ser | Gly | Tyr | Glu | Glu | Leu | Leu | Ala | Asp | Ile | Val | Asn | Leu | | | | | | | |
| | 370 | | | | | 375 | | | | | 380 | | | | | | | | | | | |
| Cys | Val | Asp | Tyr | Tyr | Glu | Asn | Arg | Met | Tyr | Leu | Thr | Pro | Ser | Glu | Lys | | | | | | | |
| 385 | | | | | 390 | | | | | 395 | | | | | 400 | | | | | | | |
| His | Met | Leu | Leu | Lys | Val | Met | Gly | Phe | Gly | Leu | Tyr | Leu | Met | Asp | Gly | | | | | | | |
| | | | | 405 | | | | | 410 | | | | | | 415 | | | | | | | |
| Ser | Val | Ser | Asn | Ile | Tyr | Lys | Leu | Asp | Ala | Lys | Lys | Arg | Ile | Asn | Leu | | | | | | | |
| | | | 420 | | | | | 425 | | | | | 430 | | | | | | | | | |
| Ser | Lys | Ile | Asp | Lys | Tyr | Phe | Lys | Gln | Leu | Gln | Val | Val | Pro | Leu | Phe | | | | | | | |
| | | 435 | | | | | 440 | | | | | 445 | | | | | | | | | | |
| Gly | Asp | Met | Gln | Ile | Glu | Leu | Ala | Arg | Tyr | Ile | Lys | Thr | Ser | Ala | His | | | | | | | |
| | 450 | | | | | 455 | | | | | 460 | | | | </ | | | | | | | |

| | | | | | | |
|-------------------------|-----------------------------|-----------------------------|--|-----|--|-----|
| 545 | | 550 | | 555 | | 560 |
| Asp Lys Tyr Ser | Asn Lys Asp Cys Pro | Asp Ser Ala Glu Glu Tyr Glu | | | | |
| | 565 | 570 | | 575 | | |
| Arg Ala Thr Arg | Tyr Asn Tyr Thr Ser | Glu Glu Lys Phe Ala Leu Val | | | | |
| | 580 | 585 | | 590 | | |
| Glu Val Ile Ala Met | Ile Lys Gly Leu Gln Val | Leu Met Gly Arg Met | | | | |
| | 595 | 600 | | 605 | | |
| Glu Ser Val Phe Asn His | Ala Ile Arg His Thr | Val Tyr Ala Ala Leu | | | | |
| | 610 | 615 | | 620 | | |
| Gln Asp Phe Ser Gln Val | Thr Leu Arg Glu Pro | Leu Arg Gln Ala Ile | | | | |
| 625 | 630 | 635 | | 640 | | |
| Lys Lys Lys Lys Asn Val | Ile Gln Ser Val | Leu Gln Ala Ile Arg Lys | | | | |
| | 645 | 650 | | 655 | | |
| Thr Val Cys Asp Trp | Glu Thr Gly His Glu Pro | Phe Asn Asp Pro Ala | | | | |
| | 660 | 665 | | 670 | | |
| Leu Arg Gly Glu Lys Asp | Pro Lys Ser Gly Phe Asp | Ile Lys Val Pro | | | | |
| | 675 | 680 | | 685 | | |
| Arg Arg Ala Val Gly | Pro Ser Ser Thr Gln | Leu Tyr Met Val Arg Thr | | | | |
| | 690 | 695 | | 700 | | |
| Met Leu Glu Ser Leu Ile | Ala Asp Lys Ser Gly Ser | Lys Lys Thr Leu | | | | |
| 705 | 710 | 715 | | 720 | | |
| Arg Ser Ser Leu Glu Gly | Pro Thr Ile Leu Asp | Ile Glu Lys Phe His | | | | |
| | 725 | 730 | | 735 | | |
| Arg Glu Ser Phe Tyr Thr | His Leu Ile Asn Phe Ser | Glu Thr Leu | | | | |
| | 740 | 745 | | 750 | | |
| Gln Gln Cys Cys Asp | Leu Ser Gln Leu Trp Phe Arg | Glu Phe Phe Leu | | | | |
| | 755 | 760 | | 765 | | |
| Glu Leu Thr Met Gly Arg | Arg Ile Gln Phe Pro Ile | Glu Met Ser Met | | | | |
| | 770 | 775 | | 780 | | |
| Pro Trp Ile Leu Thr Asp | His Ile Leu Glu Thr Lys | Glu Ala Ser Met | | | | |
| 785 | 790 | 795 | | 800 | | |
| Met Glu Tyr Val Leu Tyr | Ser Leu Asp Leu Tyr Asn Asp | Ser Ala His | | | | |
| | 805 | 810 | | 815 | | |
| Tyr Ala Leu Thr Arg Phe | Asn Lys Gln Phe Leu Tyr Asp | Glu Ile Glu | | | | |
| | 820 | 825 | | 830 | | |
| Ala Glu Val Asn Leu Cys | Phe Asp Gln Phe Val Tyr Lys | Leu Ala Asp | | | | |
| | 835 | 840 | | 845 | | |
| Gln Ile Phe Ala Tyr Tyr | Lys Val Met Ala Gly Ser | Leu Leu Leu Asp | | | | |
| | 850 | 855 | | 860 | | |
| Lys Arg Leu Arg Ser Glu | Cys Lys Asn Gln Gly Ala Thr | Ile His Leu | | | | |
| 865 | 870 | 875 | | 880 | | |
| Pro Pro Ser Asn Arg Tyr | Glu Thr Leu Leu Lys Gln Arg | His Val Gln | | | | |
| | 885 | 890 | | 895 | | |
| Leu Leu Gly Arg Ser Ile | Asp Leu Asn Arg Leu Ile Thr | Gln Arg Val | | | | |
| | 900 | 905 | | 910 | | |
| Ser Ala Ala Met Tyr Lys | Ser Leu Glu Leu Ala Ile Gly | Arg Phe Glu | | | | |
| | 915 | 920 | | 925 | | |
| Ser Glu Asp Leu Thr Ser | Ile Val Glu Leu Asp Gly | Leu Leu Glu Ile | | | | |
| | 930 | 935 | | 940 | | |
| Asn Arg Met Thr His Lys | Leu Leu Ser Arg Tyr Leu Thr | Leu Asp Gly | | | | |
| 945 | 950 | 955 | | 960 | | |
| Phe Asp Ala Met Phe Arg | Glu Ala Asn His Asn Val Ser | Ala Pro Tyr | | | | |
| | 965 | 970 | | 975 | | |
| Gly Arg Ile Thr Leu His | Val Phe Trp Glu Leu Asn Tyr | Asp Phe Leu | | | | |

| | | |
|---|------|------|
| 980 | 985 | 990 |
| Pro Asn Tyr Cys Tyr Asn Gly Ser Thr Asn Arg Phe Val Arg Thr Val | | |
| 995 | 1000 | 1005 |
| Leu Pro Phe Ser Gln Glu Phe Gln Arg Asp Lys Gln Pro Asn Ala Gln | | |
| 1010 | 1015 | 1020 |
| Pro Gln Tyr Leu His Gly Ser Lys Ala Leu Asn Leu Ala Tyr Ser Ser | | |
| 1025 | 1030 | 1035 |
| Ile Tyr Gly Ser Tyr Arg Asn Phe Val Gly Pro Pro His Phe Gln Val | | |
| 1045 | 1050 | 1055 |
| Ile Cys Arg Leu Leu Gly Tyr Gln Gly Ile Ala Val Val Met Glu Glu | | |
| 1060 | 1065 | 1070 |
| Leu Leu Lys Val Val Lys Ser Leu Leu Gln Gly Thr Ile Leu Gln Tyr | | |
| 1075 | 1080 | 1085 |
| Val Lys Thr Leu Met Glu Val Met Pro Lys Ile Cys Arg Leu Pro Arg | | |
| 1090 | 1095 | 1100 |
| His Glu Tyr Gly Ser Pro Gly Ile Leu Glu Phe Phe His His Gln Leu | | |
| 1105 | 1110 | 1115 |
| Lys Asp Ile Val Glu Tyr Ala Glu Leu Lys Thr Val Cys Phe Gln Asn | | |
| 1125 | 1130 | 1135 |
| Leu Arg Glu Val Gly Asn Ala Ile Leu Phe Cys Leu Leu Ile Glu Gln | | |
| 1140 | 1145 | 1150 |
| Ser Leu Ser Leu Glu Glu Val Cys Asp Leu Leu His Ala Ala Pro Phe | | |
| 1155 | 1160 | 1165 |
| Gln Asn Ile Leu Pro Arg Val His Val Lys Glu Gly Glu Arg Leu Asp | | |
| 1170 | 1175 | 1180 |
| Ala Lys Met Lys Arg Leu Glu Ser Lys Tyr Ala Pro Leu His Leu Val | | |
| 1185 | 1190 | 1195 |
| Pro Leu Ile Glu Arg Leu Gly Thr Pro Gln Gln Ile Ala Ile Ala Arg | | |
| 1205 | 1210 | 1215 |
| Glu Gly Asp Leu Leu Thr Lys Glu Arg Leu Cys Cys Gly Leu Ser Met | | |
| 1220 | 1225 | 1230 |
| Phe Glu Val Ile Leu Thr Arg Ile Arg Ser Phe Leu Asp Asp Pro Ile | | |
| 1235 | 1240 | 1245 |
| Trp Arg Gly Pro Leu Pro Ser Asn Gly Val Met His Val Asp Glu Cys | | |
| 1250 | 1255 | 1260 |
| Val Glu Phe His Arg Leu Trp Ser Ala Met Gln Phe Val Tyr Cys Ile | | |
| 1265 | 1270 | 1275 |
| Pro Val Gly Thr His Glu Phe Thr Val Glu Gln Cys Phe Gly Asp Gly | | |
| 1285 | 1290 | 1295 |
| Leu His Trp Ala Gly Cys Met Ile Ile Val Leu Leu Gly Gln Gln Arg | | |
| 1300 | 1305 | 1310 |
| Arg Phe Ala Val Leu Asp Phe Cys Tyr His Leu Leu Lys Val Gln Lys | | |
| 1315 | 1320 | 1325 |
| His Asp Gly Lys Asp Glu Ile Ile Lys Asn Val Pro Leu Lys Lys Met | | |
| 1330 | 1335 | 1340 |
| Val Glu Arg Ile Arg Lys Phe Gln Ile Leu Asn Asp Glu Ile Ile Thr | | |
| 1345 | 1350 | 1355 |
| Ile Leu Asp Lys Tyr Leu Lys Ser Gly Asp Gly Glu Gly Thr Pro Val | | |
| 1365 | 1370 | 1375 |
| Glu His Val Arg Cys Phe Gln Pro Pro Ile His Gln Ser Leu Ala Ser | | |
| 1380 | 1385 | 1390 |
| Ser | | |

<210> 177
 <211> 417
 <212> DNA
 <213> Homo sapiens

<400> 177
 acgcgtgatg tcacactgcc tctgccgctg ggtcctaatt cgattgcacg caccatggct
 60
 gcagttcgtg gcgcgcatag tttctggcat gcttcgcgca tcctggagac cgatcccgcc
 120
 gctgccgtga aaccgcctaa aaatgtgaag cgattgccca aagccgtgtc cgtggagcaa
 180
 atgcaaaagc tccttgccat acccagtctt aagactccta ccggcctgcg taatcgagcg
 240
 atacttgagt tcttatatgc taccggcgcg cgcgtgagcg agatgctggc aacagacctg
 300
 gacgatatac acctggggcga aaaaccccgcg gatgaaaacg gggaatctat tgcacttccc
 360
 gggatatgtgc gccttttttg aaagggaggt aaagagcggt tagtcccttt gggatcc
 417

<210> 178
 <211> 139
 <212> PRT
 <213> Homo sapiens

<400> 178
 Thr Arg Asp Val Thr Leu Pro Leu Pro Leu Gly Pro Asn Ser Ile Ala
 1 5 10 15
 Arg Thr Met Ala Ala Val Arg Gly Ala His Ser Phe Trp His Ala Ser
 20 25 30
 Arg Ile Leu Glu Thr Asp Pro Ala Ala Val Lys Pro Pro Lys Asn
 35 40 45
 Val Lys Arg Leu Pro Lys Ala Val Ser Val Glu Gln Met Gln Lys Leu
 50 55 60
 Leu Ala Ile Pro Ser Leu Lys Thr Pro Thr Gly Leu Arg Asn Arg Ala
 65 70 75 80
 Ile Leu Glu Phe Leu Tyr Ala Thr Gly Ala Arg Val Ser Glu Met Leu
 85 90 95
 Ala Thr Asp Leu Asp Asp Ile His Leu Gly Glu Lys Pro Arg Asp Glu
 100 105 110
 Asn Gly Glu Ser Ile Ala Leu Pro Gly Tyr Val Arg Leu Phe Gly Lys
 115 120 125
 Gly Gly Lys Glu Arg Leu Val Pro Leu Gly Ser
 130 135

<210> 179
 <211> 362
 <212> DNA
 <213> Homo sapiens

<400> 179
 acgcgtcgaa ggtgccggtg ggggcatca ataacatcgc gcaatccctg gaagagcctc
 60

aggtgattgc ccgtgggttg atggtggaag atccccgcatc cccaagaatc cgggaattcg
 120
 ccattggggcc gggcagcccg aatccaaaat gtcggggcac gccagtgagg agtatggtaa
 180
 ggggccggca ccgatgttgg nggcagcata cggatggaag tgctgggcga ggcctgggt
 240
 ttgccggcag agcaactggg gcagctcaag gcgggcgggg tgatcgagca gttggattga
 300
 gcaatggcgg ccgcgaagcc cgccatttac cttgatgact gtttagcgcg cggattcttt
 360
 aa
 362

<210> 180
 <211> 108
 <212> PRT
 <213> Homo sapiens

<400> 180
 Met Ala Gly Phe Ala Ala Ile Ala Gln Ser Asn Cys Ser Ile Thr
 1 5 10 15
 Pro Pro Ala Leu Ser Cys Pro Ser Cys Ser Ala Gly Lys Pro Arg Arg
 20 25 30
 Ser Pro Ser Thr Ser Ile Arg Met Leu Pro Pro Thr Ser Val Pro Ala
 35 40 45
 Pro Tyr His Thr Pro Thr Gly Arg Ala Pro Thr Phe Trp Ile Arg Ala
 50 55 60
 Ala Arg Pro Asn Gly Glu Phe Pro Asp Ser Trp Gly Cys Gly Ile Phe
 65 70 75 80
 His His Gln Pro Thr Gly Asn His Leu Arg Leu Phe Gln Gly Leu Arg
 85 90 95
 Asp Val Ile Asp Arg Pro His Arg His Leu Arg Arg
 100 105

<210> 181
 <211> 297
 <212> DNA
 <213> Homo sapiens

<400> 181
 gcgttgatca tgtccgaccc aggcttgatc atgctggtac gccgtcactt cccgtgcatg
 60
 ccgattcact tgtcgggtaca ggccaatacg gtgaattggg ccagcgtcga gttctggcaa
 120
 cagcaaggta tctgccgggt aatcctgtcg cgggaattgt cactggaaga aatcggcgaa
 180
 atccgccaac aggtgccggc catggagctg gaagtgtttg tgcacggtgc cctgtacatg
 240
 gcctattccg ggcgtgtttt gttgtccggc tatatgaaca agcgcgatgc caaccaa
 297

<210> 182
 <211> 99
 <212> PRT

<213> Homo sapiens

<400> 182

Ala Leu Ile Met Ser Asp Pro Gly Leu Ile Met Leu Val Arg Arg His
 1 5 10 15
 Phe Pro Cys Met Pro Ile His Leu Ser Val Gln Ala Asn Thr Val Asn
 20 25 30
 Trp Ala Ser Val Glu Phe Trp Gln Gln Gly Ile Cys Arg Val Ile
 35 40 45
 Leu Ser Arg Glu Leu Ser Leu Glu Glu Ile Gly Glu Ile Arg Gln Gln
 50 55 60
 Val Pro Ala Met Glu Leu Glu Val Phe Val His Gly Ala Leu Tyr Met
 65 70 75 80
 Ala Tyr Ser Gly Arg Cys Leu Leu Ser Gly Tyr Met Asn Lys Arg Asp
 85 90 95
 Ala Asn Gln

<210> 183

<211> 351

<212> DNA

<213> Homo sapiens

<400> 183

cgggagctca ccatgaagcc gaccggctcg ggggatgtgg cgaacaaggt catcacccat
 60
 attcogttta acatcgtctc ccaggcgact catccattcc ttcgtacctt ggacgatgtc
 120
 aagcgcacat ctttggcgac cgacgggctc ggccaccagg tcctgctcaa gggctaccag
 180
 gccgagggcc acgactacgc acaccccgac tacggcgga acgtctccca ccgtgccggc
 240
 gggatgaagg atctcgagaa gctcaccgag tcgggcaggc agtggaacac cgatttcggc
 300
 attcacgtca acctggtgga gtcctatcct gaggcgaatc acttcggcga c
 351

<210> 184

<211> 117

<212> PRT

<213> Homo sapiens

<400> 184

Arg Asp Val Thr Met Lys Pro Thr Gly Ser Gly Asp Val Ala Asn Lys
 1 5 10 15
 Val Ile Thr His Ile Pro Phe Asn Ile Val Ser Gln Ala Thr His Pro
 20 25 30
 Phe Leu Arg Thr Leu Asp Asp Val Lys Arg Ile Ser Leu Ala Thr Asp
 35 40 45
 Gly Leu Gly His Gln Val Leu Leu Lys Gly Tyr Gln Ala Glu Gly His
 50 55 60
 Asp Tyr Ala His Pro Asp Tyr Gly Gly Asn Val Ser His Arg Ala Gly
 65 70 75 80
 Gly Met Lys Asp Leu Glu Lys Leu Thr Glu Ser Gly Arg Gln Trp Asn

```

      85          90          95
Thr Asp Phe Gly Ile His Val Asn Leu Val Glu Ser Tyr Pro Glu Ala
      100          105          110
Asn His Phe Gly Asp
      115

```

<210> 185
 <211> 396
 <212> DNA
 <213> Homo sapiens

```

<400> 185
cgcggtgggtc tcagtaaaga aaatttggtg cttagaggat gcaccattag aaacacagag
60
gctgttggtg gcattgtggt ttatgcaggc catgaaacca aagcaatgct gaacaacagt
120
gggccacggt ataagcgcag caaattagaa agaagagcaa acacagatgt cctctggtgt
180
gtcatgcttc tggtcataat gtgcttaact ggcgcagtag gtcattggaat ctggctgagc
240
aggatatgaaa agatgcattt tttcaatggt cccgagcctg atggacatat catatcacca
300
ctgttggcag gattttatat gttttggacc gtgatcattt tgttacaggt cttgattcct
360
atttctctct atgtttccat cgaaattgtg aagctt
396

```

<210> 186
 <211> 132
 <212> PRT
 <213> Homo sapiens

```

<400> 186
Arg Val Gly Leu Ser Lys Glu Asn Leu Leu Leu Arg Gly Cys Thr Ile
1      5      10      15
Arg Asn Thr Glu Ala Val Val Gly Ile Val Val Tyr Ala Gly His Glu
20     25     30
Thr Lys Ala Met Leu Asn Asn Ser Gly Pro Arg Tyr Lys Arg Ser Lys
35     40     45
Leu Glu Arg Arg Ala Asn Thr Asp Val Leu Trp Cys Val Met Leu Leu
50     55     60
Val Ile Met Cys Leu Thr Gly Ala Val Gly His Gly Ile Trp Leu Ser
65     70     75     80
Arg Tyr Glu Lys Met His Phe Phe Asn Val Pro Glu Pro Asp Gly His
85     90     95
Ile Ile Ser Pro Leu Leu Ala Gly Phe Tyr Met Phe Trp Thr Val Ile
100    105    110
Ile Leu Leu Gln Val Leu Ile Pro Ile Ser Leu Tyr Val Ser Ile Glu
115    120    125
Ile Val Lys Leu
130

```

<210> 187
 <211> 423

<212> DNA

<213> Homo sapiens

<400> 187

cgagtgtctca ccgcgtcagc cgtcatgcgt ccactgagg ctgttgcttc tcggtcggca
 60
 gaacctcgac gagttcagcg gatcctggac cagcgcgagt gggctggcgt ctctgttgct
 120
 gatgagcatc gtcgtttgct tggcacggtc ggcgatcaag aggtcatcga ggctgctcgc
 180
 cgcggagatc gcagtattgc tgacgcgggtg gaaactaacg gcacccctcac ggcgcggacc
 240
 gacactccgt tgtccgagct ctctgctccg accagcaacg ccagggtgcc gttggccgtt
 300
 gtcgacgagg acttccacct catgggtgtc atctctcggg tgacctgct cgacgcgatg
 360
 tcacgagctc gcgacgaggc aggagaggga tctgtcatgt ccttgagaaa caccggaaag
 420
 ctt
 423

<210> 188

<211> 141

<212> PRT

<213> Homo sapiens

<400> 188

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Arg | Val | Leu | Thr | Ala | Ser | Ala | Val | Met | Arg | Pro | Thr | Glu | Ala | Val | Val |
| 1 | | | | 5 | | | | 10 | | | | | 15 | | |
| Ser | Arg | Ser | Ala | Glu | Pro | Arg | Arg | Val | Gln | Arg | Ile | Leu | Asp | Gln | Arg |
| | | 20 | | | | | 25 | | | | | 30 | | | |
| Glu | Trp | Ala | Gly | Val | Phe | Val | Val | Asp | Glu | His | Arg | Arg | Leu | Leu | Gly |
| | 35 | | | | | 40 | | | | | 45 | | | | |
| Thr | Val | Gly | Asp | Gln | Glu | Val | Ile | Glu | Ala | Ala | Arg | Arg | Gly | Asp | Arg |
| | 50 | | | | | 55 | | | | | 60 | | | | |
| Ser | Ile | Ala | Asp | Ala | Val | Glu | Thr | Asn | Gly | Ile | Leu | Thr | Ala | Arg | Thr |
| | 65 | | | | 70 | | | | 75 | | | | | 80 | |
| Asp | Thr | Pro | Leu | Ser | Glu | Leu | Phe | Ala | Pro | Thr | Ser | Asn | Ala | Arg | Val |
| | | | 85 | | | | | 90 | | | | | 95 | | |
| Pro | Leu | Ala | Val | Val | Asp | Glu | Asp | Phe | His | Leu | Met | Gly | Val | Ile | Ser |
| | | 100 | | | | | 105 | | | | | 110 | | | |
| Arg | Val | Thr | Leu | Leu | Asp | Ala | Met | Ser | Arg | Ala | Arg | Asp | Glu | Ala | Gly |
| | | 115 | | | | | 120 | | | | | 125 | | | |
| Glu | Gly | Ser | Val | Met | Ser | Leu | Glu | Asn | Thr | Gly | Lys | Leu | | | |
| | 130 | | | | | 135 | | | | | | 140 | | | |

<210> 189

<211> 429

<212> DNA

<213> Homo sapiens

<400> 189

ngatgggttta ccaacatatg cacgggttcga gcggcaatag ctccctcgggg gctggcagtg
 60

aaatgtttga agatgccggc gtttcgggcc tcaacttggt tcgatgccgt gggtccaccg
 120
 atttcgccga tgcggctcat cgcacgggta agaagtttcg tccagataac ccaggacaga
 180
 gcaaggtata tcaggctcag aaccaggaaa agcagggcct taccacagtg ccccatatag
 240
 accgcgctag ctacggcaaa aggcgcgccc agtgggggtcc aggacagcac tttcatggct
 300
 gaaggagcg catcccnagc ttcgcctagc cccagagcta acccagcgac cagtggacca
 360
 gcgccccatca tcagtaggaa ccctacgata atcagccctt gttttacccc tggaatggag
 420
 ctgatttcn
 429

<210> 190

<211> 123

<212> PRT

<213> Homo sapiens

<400> 190

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Met | Gly | Ala | Gly | Pro | Leu | Val | Ala | Gly | Leu | Ala | Leu | Gly | Leu | Gly |
| 1 | | | | 5 | | | | | 10 | | | | | 15 | |
| Glu | Ala | Xaa | Asp | Ala | Leu | Pro | Ser | Ala | Met | Lys | Val | Leu | Ser | Trp | Thr |
| | | | 20 | | | | | 25 | | | | | 30 | | |
| Pro | Leu | Gly | Ala | Pro | Phe | Ala | Val | Ala | Ser | Ala | Val | Tyr | Met | Gly | His |
| | | 35 | | | | 40 | | | | | 45 | | | | |
| Trp | Gly | Lys | Ala | Leu | Leu | Phe | Leu | Val | Leu | Ser | Leu | Ile | Tyr | Leu | Ala |
| | 50 | | | | 55 | | | | | 60 | | | | | |
| Leu | Ser | Trp | Val | Ile | Trp | Thr | Lys | Leu | Leu | Asn | Arg | Ala | Met | Ser | Arg |
| 65 | | | | 70 | | | | | 75 | | | | | 80 | |
| Ile | Gly | Glu | Ile | Gly | Gly | Thr | Thr | Ala | Ser | Lys | Gln | Val | Glu | Ala | Gly |
| | | | 85 | | | | | 90 | | | | | 95 | | |
| Asn | Ala | Gly | Ile | Phe | Lys | His | Phe | Thr | Ala | Ser | Pro | Arg | Gly | Ala | Ile |
| | | 100 | | | | | 105 | | | | | | 110 | | |
| Ala | Ala | Arg | Thr | Val | His | Met | Leu | Val | Asn | His | | | | | |
| | | 115 | | | | | 120 | | | | | | | | |

<210> 191

<211> 4845

<212> DNA

<213> Homo sapiens

<400> 191

ccgccccggg ccatggcgac actcagcttc gtcttctctgc tgctgggggc agtgtctctg
 60
 cctccggctt ctgcctccgg ccaggagtgc tggccccggac aatcggcggc cgatattctg
 120
 tcgggggagg cttcccgcag acggtatctt ctgtatgacg tcaaccccc ggaaggcttc
 180
 aacctgcgca gggatgtcta tatccgaatc gcctctctcc tgaagactct gctgaagacg
 240
 gaggagtggg tgcttgtcct gcctccatgg ggccgcctct atcactggca gagtctctgac
 300

atccaccagg tccggattcc ctggtctgag ttttttgatc ttccaagtct caataaaaaac
360
atccccgtca tcgagtatga gcagttcatc gcagaatctg gtggggccctt tattgaccag
420
gtttacgtcc tgcaaagtta cgcagagggg tggaagaag ggacctggga agagaagggtg
480
gacgagcggc cgtgtattga tcagctcctg tactcccagg acaagcacga gtactacaga
540
ggatgggtttt ggggttatga ggagaccagg ggtctaaacg tctcctgtct gtccgtccag
600
ggctcagcct ccatcgtggc gccctgctg ctgagaaaca catcagcccg gtccgtgatg
660
ttagacagag ccgagaacct acttcacgac cactatggag ggaaagaata ctgggatacc
720
cgtcgcagca tgggtgtttgc caggcacctg cgggaggtgg gagacgagtt caggagcaga
780
catctcaact ccacggacga cgcagacagg atccccctcc aggaggactg gatgaagatg
840
aagggtcaagc tgggctccgc gctagggggc ccctacctgg gagtccacct gagaagaaaa
900
gatttcatct ggggtcacag acaggatgta cccagtctgg aaggggccgt gaggaagatc
960
cgcagcctca tgaagacca ccggctggac aagggtgtttg tggccacaga tgccgtcaga
1020
aaggaatatg aagagctaaa aaagctgtta cccgagatgg tgaggtttga acccacgtgg
1080
gaggagctgg agctctacaa ggacggaggc gttgcgatta ttgaccagtg gatctgcgca
1140
cacgccaggt gcctgcccac gtcactgtcg gccgagagcg ggtcgggtgg ctttcaaagg
1200
ttcttctgtc ccaagtactc ggtgtcagag cagatggctg cctgtgttca cagtggcat
1260
ttccatactg ttgacctc cgtctgagtc tcctgtagca tctggttcag tgtttccctg
1320
ggctgaagtt aattgttcat cttgccccct tagttctcat gcacagaatt cctccatagc
1380
aggctgttgg catagctggc ctctgtctcag aacctcttct tgtgtcgcat tttcccatca
1440
ttcccggttt ctgccccctg ctgccccctg ccctgagagt tggccgtgcc ctggacttgg
1500
gcatgtcctt gttgctgtgt tgttgagcat ccgtgagcgt ccccgaggcc gggagcgtgg
1560
gccctcgtgt gatcattctc gtggggctgc catgagcgtc cccaaggctg ggagcatggg
1620
ccctcgtgtg atcgttcttg tggggctgcc gtgagcgtcc ccgaggccgg gagcgtgggg
1680
cctcgtgtga tcattctcgt ggggctgccg tgagcgtccc agaggccggg agcgtggggc
1740
ctgctgcag tcattcctct ggggttctg tgggaggtac gcctgggct ctgttcctcc
1800
aaagacctgc ctgcccctc gcataaggaga tgaaggctgg ggtaggggtg aaacgggttg
1860
agttaaatgg aaaatgaaag tagaggggaat gatcttcccc gtgggttagca ctgtgcacac
1920

gcgtgcgtct ctgtgggtta gtctgtctct ctctgcccg aggaatgctg agcgccctga
1980
gccggtgcct cttcacacat ctgctatttc ctgtggtgtt ctgggcatgg tgtataagac
2040
ccacagagge tccgggtgat gctgtctgct ggggtgtgggt ccccttcctt gttaagcaga
2100
caggatgcag cgctgacttc ttaggtcagg gcggaggtgt gcaggagccc agtcacgagc
2160
tcacccctgc ttctcaggtg tggccttggg attttgactg cgaccttggc ggtgctgtct
2220
ccgcagccca ggaagcctgc tgtggggagg ctctgcactg agctctcagc ctctgccc
2280
cagctgcgcg aagcgctcgg cccagctcac tgaagctgcc ctgcctccgg gccggcgcg
2340
cctgctctgg caggccctg tgtgtgggggt ggtgagggtc tccccaccag tgctgcaccc
2400
cgcagcagca tacaggcctg tgtggcctgc tggccctgtg gctctgtgta cagcgctgtg
2460
catgttacat ttgctctgga aacatctctg gggtttgctt gttcacgaag ttcataagt
2520
gccgctggag agccagagac cagctgcgca ggagccggag gaacgggcag gccgctgacc
2580
tgaggctctg agaaaccctt ggagaagggt gtccccacca gccatacag cgtgtgtgtg
2640
gagggggcct tgacctcgt gatgtctact gtgcctcagg ataaggaccc gccatgccct
2700
ggctagacag tgtgtggtta gtaggaatct ctcatgttc accatgtgac cccaggagg
2760
tattcgacct gcactggcgt gcctggcctg ggatttggtg acggagagga gggctcccag
2820
ggggacatgc ggtgggacag gagcgtggcg gctgctgact gtggggtgtg gatggggctg
2880
cagcaccagg cagagccctt cagggcgccg atgtcgaggg cacctgagcg aggggtgcca
2940
gcaagggggc ggccggctgg tgggtgctcg ggggacggcc gtgttggttc catgtaactg
3000
tggaactcat ttcactcagc tgctcctctc agttctccct gactctggaa cctctgtgac
3060
cccagttctc cctgactctg gaacctctgt gatcctttgc aggttttttt attggcacct
3120
cagtctcaac attttctttt cggattcatg aggaagaga aatcctgggg ttggacccca
3180
agacgacgta caacagggtt tgcggagacc aagagaaggc gtgtgagcaa cccaccact
3240
ggaagatcac ctactgagga ggatcctcca gggccgctcc ccggaccoga caggcgcggg
3300
tggtatgcagg ttctgtcgcc gtggagtcac cgtctactgc cagccgggag ctgggaggac
3360
aggaccgtcc ctgcgagggt cccaggccca gaagaggccc cagcctcta gagctgggct
3420
ccgtcctcgg cgttgccagc cgccatggct gatgaagagg ctccgctgct ctggggggcg
3480
gcggttgttt tcaggcagcg tctgtgaacc cacagctcgg ttgccagcag tgcccgcg
3540

gtgaccacaga agcaggagtg tttgtcaggc tcccgtctcg gcctttccag ccacctttca
 3600
 tgtcttcata ttttaagtgc attgaggata gatgcaggcg ggtgagctgc cctccgtcag
 3660
 gtggaccggt gctgacattt ccctgggagc tgggtgcaagg agaagcgtca ttttaaagt
 3720
 ctgcagagcg accagggggc tcatgaatct ctccgttgcc ctccgcgcag caggaggctg
 3780
 cctgtgtgtt tcttcctggg acgcgtgcaa ggcagacctg gtgctgcaaa ggaaagggc
 3840
 tgaggcctca gggagccccg tggagggtatg acagttcagg ccctactgct ggcacgtcag
 3900
 agcactggga agtttttcag tgacgtctct ggggcactca gtggattgtc tgtaggaaac
 3960
 ttgcagctct gctcctcaca ccaggccccg ctggccaccc accctcgccc ccactggcca
 4020
 cccctccctc gccccgactg ccccgcccca cctcaccctc gactgccccg ccctcgcccc
 4080
 gctggccgtc cctgccctcg ccccggttgg cagggtgcaca tggggcctcc aggtctgcca
 4140
 ttgctattg agaactagaa atgaggaagg acagttacgc taactccaaa aggtgtctta
 4200
 ggatgagctg ctttatcagg gagctccttg taccattttt acagaaatca ttttaggtc
 4260
 tttgtgccac caccacgagg ggcattctga aagagggcaa cgctagacac agaatccgtg
 4320
 gaaggtgcag cagtgcctca ggggtcctca gggtcaggga gccccctca ccctcttggc
 4380
 ccgttacctt ttgtgacttt ccaccatggt gtcgtgtgac cctcagtcag gttggtggg
 4440
 gctgagtcct cactgagcag ccactttcca catctgctag aggaacagtg acatggacac
 4500
 ctgtgacaga gagaggacag ttagtgagga gggacagaca gctcttcctt tcggagcctg
 4560
 gctagtctag gacatcacct tgctgtgtct tctcaagctt ttaaaattga ccctgaacgt
 4620 cctatggtgt tactcaaagc tgtgcagggt aaatgatgac atatttatc 4680
 tttttccatt tgttctagaa acagtgcctt tttcatcagt tgcattttcc aggtgagag
 4740
 ctgtataaaa cattttggac tgtgaccatg taccttcctt ttttaagaaaa ataaactgct
 4800
 ttatggaagt tggtaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaa
 4845

<210> 192

<211> 428

<212> PRT

<213> Homo sapiens

<400> 192

Pro Pro Gly Ala Met Ala Thr Leu Ser Phe Val Phe Leu Leu Leu Gly
 1 5 10 15
 Ala Val Ser Trp Pro Pro Ala Ser Ala Ser Gly Gln Glu Phe Trp Pro
 20 25 30
 Gly Gln Ser Ala Ala Asp Ile Leu Ser Gly Ala Ala Ser Arg Arg Arg

| 35 | | | | 40 | | | | 45 | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Tyr | Leu | Leu | Tyr | Asp | Val | Asn | Pro | Pro | Glu | Gly | Phe | Asn | Leu | Arg | Arg |
| 50 | | | | | | 55 | | | | | 60 | | | | |
| Asp | Val | Tyr | Ile | Arg | Ile | Ala | Ser | Leu | Leu | Lys | Thr | Leu | Leu | Lys | Thr |
| 65 | | | | | 70 | | | | | 75 | | | | | 80 |
| Glu | Glu | Trp | Val | Leu | Val | Leu | Pro | Pro | Trp | Gly | Arg | Leu | Tyr | His | Trp |
| | | | | 85 | | | | | 90 | | | | | 95 | |
| Gln | Ser | Pro | Asp | Ile | His | Gln | Val | Arg | Ile | Pro | Trp | Ser | Glu | Phe | Phe |
| | | | | 100 | | | | 105 | | | | | 110 | | |
| Asp | Leu | Pro | Ser | Leu | Asn | Lys | Asn | Ile | Pro | Val | Ile | Glu | Tyr | Glu | Gln |
| | 115 | | | | | | 120 | | | | | 125 | | | |
| Phe | Ile | Ala | Glu | Ser | Gly | Gly | Pro | Phe | Ile | Asp | Gln | Val | Tyr | Val | Leu |
| | 130 | | | | | 135 | | | | | 140 | | | | |
| Gln | Ser | Tyr | Ala | Glu | Gly | Trp | Lys | Glu | Gly | Thr | Trp | Glu | Glu | Lys | Val |
| 145 | | | | | 150 | | | | | 155 | | | | | 160 |
| Asp | Glu | Arg | Pro | Cys | Ile | Asp | Gln | Leu | Leu | Tyr | Ser | Gln | Asp | Lys | His |
| | | | | 165 | | | | | 170 | | | | | 175 | |
| Glu | Tyr | Tyr | Arg | Gly | Trp | Phe | Trp | Gly | Tyr | Glu | Glu | Thr | Arg | Gly | Leu |
| | | | | 180 | | | | 185 | | | | | 190 | | |
| Asn | Val | Ser | Cys | Leu | Ser | Val | Gln | Gly | Ser | Ala | Ser | Ile | Val | Ala | Pro |
| | 195 | | | | | | 200 | | | | | 205 | | | |
| Leu | Leu | Leu | Arg | Asn | Thr | Ser | Ala | Arg | Ser | Val | Met | Leu | Asp | Arg | Ala |
| | 210 | | | | | 215 | | | | | 220 | | | | |
| Glu | Asn | Leu | Leu | His | Asp | His | Tyr | Gly | Gly | Lys | Glu | Tyr | Trp | Asp | Thr |
| 225 | | | | | 230 | | | | | 235 | | | | | 240 |
| Arg | Arg | Ser | Met | Val | Phe | Ala | Arg | His | Leu | Arg | Glu | Val | Gly | Asp | Glu |
| | | | | 245 | | | | | 250 | | | | | 255 | |
| Phe | Arg | Ser | Arg | His | Leu | Asn | Ser | Thr | Asp | Asp | Ala | Asp | Arg | Ile | Pro |
| | | | | 260 | | | | 265 | | | | | 270 | | |
| Phe | Gln | Glu | Asp | Trp | Met | Lys | Met | Lys | Val | Lys | Leu | Gly | Ser | Ala | Leu |
| | 275 | | | | | | 280 | | | | | 285 | | | |
| Gly | Gly | Pro | Tyr | Leu | Gly | Val | His | Leu | Arg | Arg | Lys | Asp | Phe | Ile | Trp |
| | 290 | | | | | 295 | | | | | 300 | | | | |
| Gly | His | Arg | Gln | Asp | Val | Pro | Ser | Leu | Glu | Gly | Ala | Val | Arg | Lys | Ile |
| 305 | | | | | 310 | | | | | 315 | | | | | 320 |
| Arg | Ser | Leu | Met | Lys | Thr | His | Arg | Leu | Asp | Lys | Val | Phe | Val | Ala | Thr |
| | | | | 325 | | | | | 330 | | | | | 335 | |
| Asp | Ala | Val | Arg | Lys | Glu | Tyr | Glu | Glu | Leu | Lys | Lys | Leu | Leu | Pro | Glu |
| | | | | 340 | | | | 345 | | | | | 350 | | |
| Met | Val | Arg | Phe | Glu | Pro | Thr | Trp | Glu | Glu | Leu | Glu | Leu | Tyr | Lys | Asp |
| | 355 | | | | | | 360 | | | | | 365 | | | |
| Gly | Gly | Val | Ala | Ile | Ile | Asp | Gln | Trp | | | | | | | |

```
<210> 193
<211> 350
<212> DNA
<213> Homo sapiens
```

<400> 193

gccggcgagc tggactgcgc catcatggcc gagcccttcc ccgacaccgg cctggccacg
60
gcgcagctgt acgacgagcc cttcgtcgtc gcgctgcggg cgtcgcaccc gctggccgac
120
cgtgccagca tcagccccga ggaggtcaag ggcgagacca tgttgatgtt gggcacgggc
180
ccctgggttc cccggggccc cggtgggggt ttggcccga tttggcgcgt ttctccagcg
240
ccgttaaggg catacgccgc agtttcgagg gctcgtcgtt ggagaccatc aagcacatcg
300
tggcttcggg catggcgtga cggtggtgcc gcagctgtcc gtgccgcgcg
350

<210> 194

<211> 116

<212> PRT

<213> Homo sapiens

<400> 194

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Ala | Gly | Glu | Leu | Asp | Cys | Ala | Ile | Met | Ala | Glu | Pro | Phe | Pro | Asp | Thr |
| 1 | | | | 5 | | | | | 10 | | | | | 15 | |
| Gly | Leu | Ala | Thr | Ala | Gln | Leu | Tyr | Asp | Glu | Pro | Phe | Val | Val | Ala | Leu |
| | | | 20 | | | | | 25 | | | | | 30 | | |
| Arg | Ala | Ser | His | Pro | Leu | Ala | Asp | Arg | Ala | Ser | Ile | Ser | Pro | Glu | Glu |
| | | 35 | | | | | 40 | | | | | 45 | | | |
| Val | Lys | Gly | Glu | Thr | Met | Leu | Met | Leu | Gly | Thr | Gly | Pro | Trp | Phe | Pro |
| | 50 | | | | 55 | | | | | 60 | | | | | |
| Arg | Ala | Arg | Gly | Gly | Gly | Leu | Ala | Arg | Ile | Trp | Arg | Val | Ser | Pro | Ala |
| 65 | | | | | 70 | | | | 75 | | | | | 80 | |
| Pro | Leu | Arg | Ala | Tyr | Ala | Ala | Val | Ser | Arg | Ala | Arg | Arg | Trp | Arg | Pro |
| | | | 85 | | | | | 90 | | | | | 95 | | |
| Ser | Ser | Thr | Ser | Trp | Leu | Arg | Ala | Trp | Arg | Asp | Gly | Gly | Ala | Ala | Ala |
| | | | 100 | | | | | 105 | | | | | 110 | | |
| Val | Arg | Ala | Ala | | | | | | | | | | | | |
| | | | 115 | | | | | | | | | | | | |

<210> 195

<211> 495

<212> DNA

<213> Homo sapiens

<400> 195

acgcgtgaac gcgacggctt ggcgatcgga ggcgtcggcc ccgctcgttga gtgggcccgtt
60
gaaatgggtc gcttcgacga aagcgagact ctcgaccgcc ttgcatcggg cgtccttgaa
120
ccagaacttg gcgacgattt ggccgccgtc ctgctcgatt ctcacgggt tgctgtcatc
180
agcgagggat cgaactggct tgcctcgcta cccgtgatcg taggtcgcaa cacggaacag
240
tttcgcagca taccagacct tgcccgcgac cggatcgaca aactgcacca gttgagccat
300

cgcgaaatag cacgaaatcg cgagctcctg cgtgcccgcg ctgcgtcggg gcaggtgcgg
 360
 cactgccacg gcgacgcaca cctcggcaac atcgtcatga ttgacggcaa gccggctcctg
 420
 ttgcagcga tcgaatttga tcttgatata gcgacaacgg atgtgctgta cgatttcgcg
 480
 ttccctctga tggat
 495

<210> 196
 <211> 165
 <212> PRT
 <213> Homo sapiens

<400> 196
 Thr Arg Glu Arg Asp Gly Leu Ala Ile Gly Gly Val Gly Pro Val Val
 1 5 10 15
 Glu Trp Ala Val Glu Met Val Arg Phe Asp Glu Ser Glu Thr Leu Asp
 20 25 30
 Arg Leu Ala Ser Gly Val Leu Glu Pro Glu Leu Gly Asp Asp Leu Ala
 35 40 45
 Ala Val Leu Leu Asp Ser His Arg Val Ala Val Ile Ser Glu Gly Ser
 50 55 60
 Asn Trp Leu Ala Ser Leu Pro Val Ile Val Gly Arg Asn Thr Glu Gln
 65 70 75 80
 Phe Arg Ser Ile Pro Asp Leu Ala Arg Asp Arg Ile Asp Lys Leu His
 85 90 95
 Gln Leu Ser His Arg Glu Ile Ala Arg Asn Arg Glu Leu Leu Arg Ala
 100 105 110
 Arg Ala Ala Ser Gly Gln Val Arg His Cys His Gly Asp Ala His Leu
 115 120 125
 Gly Asn Ile Val Met Ile Asp Gly Lys Pro Val Leu Phe Asp Ala Ile
 130 135 140
 Glu Phe Asp Pro Asp Ile Ala Thr Thr Asp Val Leu Tyr Asp Phe Ala
 145 150 155 160
 Phe Pro Leu Met Asp
 165

<210> 197
 <211> 402
 <212> DNA
 <213> Homo sapiens

<400> 197
 caagcaatgc ttgacgcagt tgttgaatac ttaccagcac cgactgatata tccagcaatc
 60
 aaaggtatca atccagatga aactgaaggt gaacgtcacg caagcgatga tgagccattc
 120
 tcttcattag cattcaaaat tgcaactgac ccattcgtag gtaacttaac cttcttccgt
 180
 gtgtactcag gtgtaattaa ctctggtgat acagtattaa actctgtacg tcaaaaacgt
 240
 gaacgttttg gtcgtatcgt acagatgcac gctaataaac gtgaagaaat taaagaagtt
 300

cgtgcgggcg atatcgctgc agcaatcggc ttaaaagatg taactacggg tgaaccatta
 360
 tgtgtgtcg atgcaccaat cattcttgag cgtatggaat tc
 402

<210> 198
 <211> 134
 <212> PRT
 <213> Homo sapiens

<400> 198
 Gln Ala Met Leu Asp Ala Val Val Glu Tyr Leu Pro Ala Pro Thr Asp
 1 5 10 15
 Ile Pro Ala Ile Lys Gly Ile Asn Pro Asp Glu Thr Glu Gly Glu Arg
 20 25 30
 His Ala Ser Asp Asp Glu Pro Phe Ser Ser Leu Ala Phe Lys Ile Ala
 35 40 45
 Thr Asp Pro Phe Val Gly Asn Leu Thr Phe Phe Arg Val Tyr Ser Gly
 50 55 60
 Val Ile Asn Ser Gly Asp Thr Val Leu Asn Ser Val Arg Gln Lys Arg
 65 70 75 80
 Glu Arg Phe Gly Arg Ile Val Gln Met His Ala Asn Lys Arg Glu Glu
 85 90 95
 Ile Lys Glu Val Arg Ala Gly Asp Ile Ala Ala Ala Ile Gly Leu Lys
 100 105 110
 Asp Val Thr Thr Gly Glu Pro Leu Cys Ala Val Asp Ala Pro Ile Ile
 115 120 125
 Leu Glu Arg Met Glu Phe
 130

<210> 199
 <211> 507
 <212> DNA
 <213> Homo sapiens

<400> 199
 acgcgtgaag tcgtgcatag atcgggtgtga catagagaag cctccgaccc aagctgcgta
 60
 tatcgcaaa agaccaagcg accctggacg ttctagacag aactctgcta cgaggcctga
 120
 caatagtga atccccgaga acccagctat ggaagggttt ccagatgctc gaaggcctgt
 180
 cataccagag gttagggttaa actgtatgga gactttcgag gtgaaagttg actcgccggt
 240
 aaagcctgct cctaaagagg atttagatct gatagatcta tcctcagatt caacctcggg
 300
 gctgaaaaa cactctatac tctcaacctc cgacagcgac tctcttgat ttgagcctct
 360
 tccctctctc agaatagtcg agagtgcga agaagaggag acgatgaacc aaggcgatga
 420
 cggccccctc ggtaaaaatg ctgcctcttc tccctccatc cccagccatc cctccgtcct
 480
 cagcctgagc acagctccgc ttgtaca
 507

<210> 200
 <211> 153
 <212> PRT
 <213> Homo sapiens

<400> 200
 Met Glu Gly Glu Ala Ala Phe Leu Pro Glu Gly Pro Ser Ser Pro
 1 5 10 15
 Trp Phe Ile Val Ser Ser Ser Ser Ser Leu Ser Thr Ile Leu Arg Glu
 20 25 30
 Gly Arg Gly Ser Asn Thr Arg Glu Ser Leu Ser Glu Val Glu Ser Ile
 35 40 45
 Glu Cys Phe Ser Gly Pro Glu Val Glu Ser Glu Asp Arg Ser Ile Arg
 50 55 60
 Ser Lys Ser Ser Leu Gly Ala Gly Phe Thr Gly Glu Ser Thr Phe Thr
 65 70 75 80
 Ser Lys Val Ser Ile Gln Phe Asn Leu Thr Ser Gly Met Thr Gly Leu
 85 90 95
 Arg Ala Ser Gly Asn Pro Ser Ile Ala Gly Phe Ser Gly Ile Ser Leu
 100 105 110
 Leu Ser Gly Leu Val Ala Glu Phe Cys Leu Glu Arg Pro Gly Ser Leu
 115 120 125
 Gly Leu Cys Ala Ile Tyr Ala Ala Trp Val Gly Gly Phe Ser Met Ser
 130 135 140
 His Arg Ser Met His Asp Phe Thr Arg
 145 150

<210> 201
 <211> 527
 <212> DNA
 <213> Homo sapiens

<400> 201
 gatgtggcta ttatccctgt ttcccagggtg agaaacaggg tcagtgatag agctgggatg
 60
 tgtgcctgca ggctcaccag ccagtcacct cctcaccaag gatgatgttc tccgtgggtga
 120
 gctggtcctt ggtctcctgg aactcgtggc gcacctgggc cagctgcgcc tcgaaggcat
 180
 ccttctccat ctctttggct agctgcaagt tctggagctg ctcgttgagg tctgtgatct
 240
 catccacctg ctggttgagc gtgcgcttga ggaaggccac aatctccttc ttgttattgg
 300
 ccagctgctc aaactcctgg cggaacatct tctcctgcac agccagctca tcccacttcc
 360
 gctggtaccg ggctagccgg tctccagggt ctcggtatctg gatgtggtag aactccttca
 420
 tctccttggc cagaggcggc tccacggcca ccaccggctc cttcttgccc cctttcttct
 480
 tgacttcaag ctccctgcct gccttgetca cactcttttt gggaggc
 527

<210> 202

<211> 70
 <212> PRT
 <213> Homo sapiens

<400> 202
 Gly Arg Pro Gln Ser Pro Ser Cys Tyr Trp Pro Ala Ala Gln Thr Pro
 1 5 10 15
 Gly Gly Thr Ser Ser Pro Ala Gln Pro Ala His Pro Thr Ser Ala Gly
 20 25 30
 Thr Gly Leu Ala Gly Pro Pro Gly Leu Gly Ser Gly Cys Gly Arg Thr
 35 40 45
 Pro Ser Ser Pro Trp Pro Glu Ala Ala Pro Arg Pro Pro Pro Ala Pro
 50 55 60
 Ser Cys Pro Leu Ser Ser
 65 70

<210> 203
 <211> 304
 <212> DNA
 <213> Homo sapiens

<400> 203
 ngtcaccgg tggatcatgga caacgcggcc tacgtggtct acacctcggg atccaccggc
 60
 cgacccaagg gagttgtcgt caccacacac ggactcgaca gcttcgcact cgaccagcag
 120
 cgtcgattcc acgcagatca ccactctcga accctgcact tcgccacccc cagcttcgac
 180
 ggagccgtct tcgagtacct gcaggcattc ggtgtcggag ccaccatggt gatcggtccc
 240
 accgacatct acggcgggcg cgaactggca agtctcatcc gccgcgaaca cgtcactcac
 300
 gcgt
 304

<210> 204
 <211> 101
 <212> PRT
 <213> Homo sapiens

<400> 204
 Xaa Ala Pro Val Val Met Asp Asn Ala Ala Tyr Val Val Tyr Thr Ser
 1 5 10 15
 Gly Ser Thr Gly Arg Pro Lys Gly Val Val Val Thr His Thr Gly Leu
 20 25 30
 Asp Ser Phe Ala Leu Asp Gln Gln Arg Arg Phe His Ala Asp His His
 35 40 45
 Ser Arg Thr Leu His Phe Ala Thr Pro Ser Phe Asp Gly Ala Val Phe
 50 55 60
 Glu Tyr Leu Gln Ala Phe Gly Val Gly Ala Thr Met Val Ile Val Pro
 65 70 75 80
 Thr Asp Ile Tyr Gly Gly Ala Glu Leu Ala Ser Leu Ile Arg Arg Glu
 85 90 95
 His Val Thr His Ala

100

<210> 205
 <211> 356
 <212> DNA
 <213> Homo sapiens

<400> 205
 nngaattcag caatgataac tggctcaatt gaaggtaaga caacaattga gggaattaat
 60
 gcacaattaa atacagtgtt aactttatctt tcaccacaat caaaagataa agatttaatac
 120
 atgccagatc aacaagaaga aatagatatt ctgattgcaa ccgactgtat ttcagaagga
 180
 cagaacttac aagattgtga ttacttaata aactatgaca ttcattggaa tccagttcgt
 240
 atcattcaaa gatttggacg gattgatcga attggttcga agaataaatg tgtacaatta
 300
 gttaactttt ggccagatat tacattagat gaatatattg atctaaaggg acgcgt
 356

<210> 206
 <211> 118
 <212> PRT
 <213> Homo sapiens

<400> 206
 Xaa Asn Ser Ala Met Ile Thr Gly Ser Ile Glu Gly Lys Thr Thr Ile
 1 5 10 15
 Glu Gly Ile Asn Ala Gln Leu Asn Thr Val Leu Thr Leu Phe Ser Pro
 20 25 30
 Gln Ser Lys Asp Lys Asp Leu Ile Met Pro Asp Gln Gln Glu Glu Ile
 35 40 45
 Asp Ile Leu Ile Ala Thr Asp Cys Ile Ser Glu Gly Gln Asn Leu Gln
 50 55 60
 Asp Cys Asp Tyr Leu Ile Asn Tyr Asp Ile His Trp Asn Pro Val Arg
 65 70 75 80
 Ile Ile Gln Arg Phe Gly Arg Ile Asp Arg Ile Gly Ser Lys Asn Lys
 85 90 95
 Cys Val Gln Leu Val Asn Phe Trp Pro Asp Ile Thr Leu Asp Glu Tyr
 100 105 110
 Ile Asp Leu Lys Gly Arg
 115

<210> 207
 <211> 324
 <212> DNA
 <213> Homo sapiens

<400> 207
 acgcgtgcac tgtgtgtatg catggtaacg tacacgtgtg cactgtgtgt ggtgtgcatg
 60
 catggtgtgt gcacgtgtng cactgtgtgt ggatgcatgg taatgtgcac gtgtgcactg
 120

tgtgtggtgt gtatgcatgg tgtgtgcacg tgtgcactgt gtgtgtgtgt atgcatgtgt
 180
 gtgcacatgt gcactgtgtg gtgtgtatgc atggtgtgtg cacgtgtgca ctgtgtatgc
 240
 atgngtgtgt gcatgtgtgc actgtgtatg catagtgtgc acgtgtgcac tgtgtggtgt
 300
 gtatgcatgg taatgtgcac gtgt
 324

<210> 208
 <211> 108
 <212> PRT
 <213> Homo sapiens

<400> 208
 Thr Arg Ala Leu Cys Val Cys Met Val Thr Tyr Thr Cys Ala Leu Cys
 1 5 10 15
 Val Val Cys Met His Gly Val Cys Thr Cys Xaa Thr Val Cys Gly Cys
 20 25 30
 Met Val Met Cys Thr Cys Ala Leu Cys Val Val Cys Met His Gly Val
 35 40 45
 Cys Thr Cys Ala Leu Cys Val Cys Val Cys Met Cys Val His Met Cys
 50 55 60
 Thr Val Trp Cys Val Cys Met Val Cys Ala Arg Val His Cys Val Cys
 65 70 75 80
 Met Xaa Val Cys Met Cys Ala Leu Cys Met His Ser Val His Val Cys
 85 90 95
 Thr Val Trp Cys Val Cys Met Val Met Cys Thr Cys
 100 105

<210> 209
 <211> 168
 <212> DNA
 <213> Homo sapiens

<400> 209
 nnctccagag gttatgaggt tggaagcccg gtttttttca ggtgcagaaa aggctaccat
 60
 attcaagggt ccacgactcg cacctgcctt gccaatntaa catggagtgg gatacagacc
 120
 gaatgtatac ctcatgcctg cagacagcca gaaaccccg cacacgcg
 168

<210> 210
 <211> 56
 <212> PRT
 <213> Homo sapiens

<400> 210
 Xaa Ser Arg Gly Tyr Glu Val Gly Ser Pro Val Phe Phe Arg Cys Arg
 1 5 10 15
 Lys Gly Tyr His Ile Gln Gly Ser Thr Thr Arg Thr Cys Leu Ala Asn
 20 25 30
 Leu Thr Trp Ser Gly Ile Gln Thr Glu Cys Ile Pro His Ala Cys Arg

35 40 45
 Gln Pro Glu Thr Pro Ala His Ala
 50 55

<210> 211
 <211> 354
 <212> DNA
 <213> Homo sapiens

<400> 211
 tacatgggct ttgacacagt ggtggctgaa gctgcactaa ggggtgtttgg aggcaatgtc
 60
 cagctggcag ctcagaccct tgcacaccat ggaggaagcc tcccacccga cctgcagtcc
 120
 tcaggagagg actcctcccc cacaccgtcc acatcccat ctgactctgc agggacctct
 180
 agtgcctcga cagatgaaga catggagacg gaggctgtca acgaaatcct ggaggacatt
 240
 ccggagcacg aggaggacta cctggactcc acgctggagg atgaagaagt cattattgct
 300
 gaatacttgt cctgcgttga aagtataagt tctgccngca aagaacaact gatc
 354

<210> 212
 <211> 118
 <212> PRT
 <213> Homo sapiens

<400> 212
 Tyr Met Gly Phe Asp Thr Val Val Ala Glu Ala Ala Leu Arg Val Phe
 1 5 10 15
 Gly Gly Asn Val Gln Leu Ala Ala Gln Thr Leu Ala His His Gly Gly
 20 25 30
 Ser Leu Pro Pro Asp Leu Gln Phe Ser Gly Glu Asp Ser Ser Pro Thr
 35 40 45
 Pro Ser Thr Ser Pro Ser Asp Ser Ala Gly Thr Ser Ser Ala Ser Thr
 50 55 60
 Asp Glu Asp Met Glu Thr Glu Ala Val Asn Glu Ile Leu Glu Asp Ile
 65 70 75 80
 Pro Glu His Glu Glu Asp Tyr Leu Asp Ser Thr Leu Glu Asp Glu Glu
 85 90 95
 Val Ile Ile Ala Glu Tyr Leu Ser Cys Val Glu Ser Ile Ser Ser Ala
 100 105 110
 Xaa Lys Glu Gln Leu Ile
 115

<210> 213
 <211> 669
 <212> DNA
 <213> Homo sapiens

<400> 213
 attgcccaat ctcagagtgt ccaggaaagc ctggagagcc tggtgcagtc tattggggaa
 60

gttgaacaaa acctggaagg gaaacaggtg tcatcactct catcaggagt catccaggaa
 120
 gccttagcca caaatatgaa attgaagcag gacattgctc ggcaaaagag cagcttggag
 180
 gccacccgtg agatggtgac ccgattcatg gagacagcag acagtactac agcagcagtg
 240
 ctgcagggca aactggcaga ggtgagccag cggttcgaac agctctgtct acagcagcaa
 300
 gaaaaggaga gctccctaaa gaagcttcta cccagggcag agatgtttga acacctctct
 360
 ggtaagctgc agcagttcat ggaaaacaaa agtcggatgc tggcctctgg aaatcagcca
 420
 gatcaagata ttacacattt cttccaacag atccaggagc tcaatttggga aatggaagac
 480
 caacaggaga acctagatac tcttgagcac ctggtcactg aactgagctc ttgtggcttt
 540
 gcgctggact tgtgccagca tcaggacagg gtacagaatc taagaaaaga cttcacagag
 600
 ctacagaaga cagttaaaga gagagagaaa gatgcatcat cttgccagga acagttggat
 660
 gaattccgg
 669

<210> 214

<211> 223

<212> PRT

<213> Homo sapiens

<400> 214

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Ile | Ala | Gln | Ser | Gln | Ser | Val | Gln | Glu | Ser | Leu | Glu | Ser | Leu | Leu | Gln |
| 1 | | | 5 | | | | | 10 | | | | | 15 | | |
| Ser | Ile | Gly | Glu | Val | Glu | Gln | Asn | Leu | Glu | Gly | Lys | Gln | Val | Ser | Ser |
| | | 20 | | | | | 25 | | | | | 30 | | | |
| Leu | Ser | Ser | Gly | Val | Ile | Gln | Glu | Ala | Leu | Ala | Thr | Asn | Met | Lys | Leu |
| | | 35 | | | | 40 | | | | | 45 | | | | |
| Lys | Gln | Asp | Ile | Ala | Arg | Gln | Lys | Ser | Ser | Leu | Glu | Ala | Thr | Arg | Glu |
| | 50 | | | | | 55 | | | | 60 | | | | | |
| Met | Val | Thr | Arg | Phe | Met | Glu | Thr | Ala | Asp | Ser | Thr | Thr | Ala | Ala | Val |
| 65 | | | | 70 | | | | | 75 | | | | | 80 | |
| Leu | Gln | Gly | Lys | Leu | Ala | Glu | Val | Ser | Gln | Arg | Phe | Glu | Gln | Leu | Cys |
| | | | 85 | | | | | 90 | | | | | 95 | | |
| Leu | Gln | Gln | Gln | Glu | Lys | Glu | Ser | Ser | Leu | Lys | Lys | Leu | Leu | Pro | Gln |
| | | 100 | | | | | 105 | | | | | 110 | | | |
| Ala | Glu | Met | Phe | Glu | His | Leu | Ser | Gly | Lys | Leu | Gln | Gln | Phe | Met | Glu |
| | 115 | | | | | 120 | | | | | 125 | | | | |
| Asn | Lys | Ser | Arg | Met | Leu | Ala | Ser | Gly | Asn | Gln | Pro | Asp | Gln | Asp | Ile |
| | 130 | | | | 135 | | | | 140 | | | | | | |
| Thr | His | Phe | Phe | Gln | Gln | Ile | Gln | Glu | Leu | Asn | Leu | Glu | Met | Glu | Asp |
| 145 | | | | 150 | | | | 155 | | | | | | 160 | |
| Gln | Gln | Glu | Asn | Leu | Asp | Thr | Leu | Glu | His | Leu | Val | Thr | Glu | Leu | Ser |
| | | | 165 | | | | 170 | | | | | | 175 | | |
| Ser | Cys | Gly | Phe | Ala | Leu | Asp | Leu | Cys | Gln | His | Gln | Asp | Arg | Val | Gln |
| | | 180 | | | | 185 | | | | | 190 | | | | |
| Asn | Leu | Arg | Lys | Asp | Phe | Thr | Glu | Leu | Gln | Lys | Thr | Val | Lys | Glu | Arg |

| | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| | 195 | | 200 | | 205 | | | | | | | | | |
| Glu | Lys | Asp | Ala | Ser | Ser | Cys | Gln | Glu | Gln | Leu | Asp | Glu | Phe | Arg |
| | 210 | | | | | 215 | | | | | | 220 | | |

<210> 215
 <211> 814
 <212> DNA
 <213> Homo sapiens

<400> 215
 aaatttcgta cccgctccgg cacagtacga gcccttgacg atgtgagcct ggctattaag
 60
 agagggtcca tctcagccgt tatcgggcac tccggagccg gcaaattccac cctgggttcg
 120
 ctcatcaacg gattagagac tcccacgcgt ggccgcgtct tggtagacgg caccgacgtc
 180
 tcgcagctct cggacaaaagc gatgcgcccg ctacgcgcag acatcgggat gatcttccaa
 240
 cagttcaacc tattcggctc aaggaccatc tacgacaacg ttgcctatcc actcaagctg
 300
 gctcattgga agaaagcaga cgagaagaag cgcgtcaccg aattgctgag cttcgtcggg
 360
 ttgacgagca aagcctggga ccatccagac cagctctcgg gcggacagaa acagcgggtt
 420
 ggtattgccc gagcgctagc aactaaacca tcgattttgt tggctgacga gtccacctcg
 480
 gcgctggatc cagaaacgac agctgatgtc ctatccctgc tcaagcgggt caatgcggaa
 540
 ctaggggtga cggtcgtcgt catcaccac gagatggagg tcgtccgctc gattgcccag
 600
 caggtctcgg tactagcagc tggccatctc gtcgagtctg gaagcgcgcg ccaggtcttc
 660
 gtcattccac agtcagagac caccagcgt ttctggcgca cgattatcgg ccagcaccg
 720
 agtggggagg aacaggcacg gttgcagtcg gaaaaccag atgcacgact cgtcgacgtc
 780
 agttcggtgg ccagtcactc gttcggtgac gcgt
 814

<210> 216
 <211> 271
 <212> PRT
 <213> Homo sapiens

<400> 216
 Lys Phe Arg Thr Arg Ser Gly Thr Val Arg Ala Leu Asp Asp Val Ser
 1 5 10 15
 Leu Ala Ile Lys Arg Gly Ser Ile Ser Ala Val Ile Gly His Ser Gly
 20 25 30
 Ala Gly Lys Ser Thr Leu Val Arg Leu Ile Asn Gly Leu Glu Thr Pro
 35 40 45
 Thr Arg Gly Arg Val Leu Val Asp Gly Thr Asp Val Ser Gln Leu Ser
 50 55 60
 Asp Lys Ala Met Arg Pro Leu Arg Ala Asp Ile Gly Met Ile Phe Gln

```

65          70          75          80
Gln Phe Asn Leu Phe Gly Ser Arg Thr Ile Tyr Asp Asn Val Ala Tyr
          85          90          95
Pro Leu Lys Leu Ala His Trp Lys Lys Ala Asp Glu Lys Lys Arg Val
          100          105          110
Thr Glu Leu Leu Ser Phe Val Gly Leu Thr Ser Lys Ala Trp Asp His
          115          120          125
Pro Asp Gln Leu Ser Gly Gly Gln Lys Gln Arg Val Gly Ile Ala Arg
          130          135          140
Ala Leu Ala Thr Lys Pro Ser Ile Leu Leu Ala Asp Glu Ser Thr Ser
145          150          155          160
Ala Leu Asp Pro Glu Thr Thr Ala Asp Val Leu Ser Leu Leu Lys Arg
          165          170          175
Val Asn Ala Glu Leu Gly Val Thr Val Val Val Ile Thr His Glu Met
          180          185          190
Glu Val Val Arg Ser Ile Ala Gln Gln Val Ser Val Leu Ala Ala Gly
          195          200          205
His Leu Val Glu Ser Gly Ser Ala Arg Gln Val Phe Ala His Pro Gln
          210          215          220
Ser Glu Thr Thr Gln Arg Phe Leu Ala Thr Ile Ile Gly Gln His Pro
225          230          235          240
Ser Gly Glu Glu Gln Ala Arg Leu Gln Ser Glu Asn Pro Asp Ala Arg
          245          250          255
Leu Val Asp Val Ser Ser Val Ala Ser His Ser Phe Gly Asp Ala
          260          265          270

```

<210> 217
 <211> 500
 <212> DNA
 <213> Homo sapiens

<400> 217
 nnacgcgtcg cgatgaaaga ggcgctgaaa ggtgccatcc agattccaac agtgactttt
 60
 agctctgaga agtccaatac tacagccctg gctgagttcg gaaaatacat tcataaagtc
 120
 tttcctacag tggtcagcac cagctttatc cagcatgaag tcgtggaaga gtatagccac
 180
 ctgttcacta tccaaggctc ggaccccagc ttgcagccct acctgctgat ggctcacttt
 240
 gatgtggtgc ctgcccctga agaaggctgg gaggtgcccc cattctctgg gttggagcgt
 300
 gatggcgtca tctatggttg gggcacactg gacgacaaga actctgtgat ggcattactg
 360
 caggccttgg agctcctgct gatcaggaag tacatcccc gaagatcttt cttcatttct
 420
 ctgggccatg atgaggagtc atcagggaca ggggctcaga g gatctcagc cctgctacag
 480
 tcaaggggcg tccagctagc
 500

<210> 218
 <211> 166
 <212> PRT

<213> Homo sapiens

<400> 218

```

Xaa Arg Val Ala Met Lys Glu Ala Leu Lys Gly Ala Ile Gln Ile Pro
 1           5           10           15
Thr Val Thr Phe Ser Ser Glu Lys Ser Asn Thr Thr Ala Leu Ala Glu
          20           25           30
Phe Gly Lys Tyr Ile His Lys Val Phe Pro Thr Val Val Ser Thr Ser
          35           40           45
Phe Ile Gln His Glu Val Val Glu Glu Tyr Ser His Leu Phe Thr Ile
          50           55           60
Gln Gly Ser Asp Pro Ser Leu Gln Pro Tyr Leu Leu Met Ala His Phe
65           70           75           80
Asp Val Val Pro Ala Pro Glu Glu Gly Trp Glu Val Pro Pro Phe Ser
          85           90           95
Gly Leu Glu Arg Asp Gly Val Ile Tyr Gly Trp Gly Thr Leu Asp Asp
          100          105          110
Lys Asn Ser Val Met Ala Leu Leu Gln Ala Leu Glu Leu Leu Leu Ile
          115          120          125
Arg Lys Tyr Ile Pro Arg Arg Ser Phe Phe Ile Ser Leu Gly His Asp
          130          135          140
Glu Glu Ser Ser Gly Thr Gly Ala Gln Arg Ile Ser Ala Leu Leu Gln
145          150          155          160
Ser Arg Gly Val Gln Leu
          165

```

<210> 219

<211> 361

<212> DNA

<213> Homo sapiens

<400> 219

```

acgcgttgaa acgggtatat tggggatgac gccgctgtgc aatatgcgca aggccataca
60
caagggtccgc acgctcccat gtccctcggt ttgcacagtt cttttgcgcc gcattatggc
120
gaagcgcgtc agattgcgcc tgatatcaag cgcacacagg tcaacaaccc cagccccttc
180
acttttttcg gcaccaacag ttatctgatc ggccgcgata cgctggcatt gatcgatccc
240
gggtccgcttg acgaggccca tcacgcggcg ctgctgctg ccattgccgg ccggccggtc
300
agccatatct ttgtcagcca cacacaccgg gaccactcgc cagtcgacac gggtttgaaa
360
g
361

```

<210> 220

<211> 102

<212> PRT

<213> Homo sapiens

<400> 220

```

Met Ala Asp Arg Pro Ala Gly Asn Gly Thr Gln Gln Arg Arg Val Met

```

```

      1             5             10             15
Gly Leu Val Lys Arg Thr Gly Ile Asp Gln Cys Gln Arg Ile Ala Ala
      20             25             30
Asp Gln Ile Thr Val Gly Ala Glu Lys Ser Glu Gly Ala Gly Val Val
      35             40             45
Asp Arg Asp Ala Leu Asp Ile Arg Arg Asn Leu Asp Gly Phe Ala Ile
      50             55             60
Met Arg Arg Lys Arg Thr Val Glu Asn Glu Gly His Gly Ser Val Arg
65             70             75             80
Thr Leu Cys Met Ala Leu Arg Ile Leu His Ser Gly Val Ile Pro Asn
      85             90             95
Ile Pro Val Ser Thr Arg
      100

```

<210> 221
 <211> 401
 <212> DNA
 <213> Homo sapiens

```

<400> 221
agatctctgt gtcgtcggct gcaaagagga tgagcccaga tgcatatcag gggctccctc
60
ccacatccca cctgctcggg cagcccacgg cagcccaca ctgctgcagc acacctcgt
120
gcagctctgg ttctcctca gaaatatccc tgccaccctg ctaagccttg gccaacactg
180
caccctgtcc caatgctggc ccagtgaaca cccccagg gcataccctc ctacagagca
240
ttcccaaaaa aggctagagt agacaccagg ctgctccgta gggggcctcc accccattct
300
ccaaggcctc caccaggga cgctggtga accagcatcc aggctggcc cacctccctg
360
ctcagagtcc atgttctgtg acaagggtgg caactgggat t
401

```

<210> 222
 <211> 124
 <212> PRT
 <213> Homo sapiens

```

<400> 222
Met Asp Ser Glu Gln Gly Gly Gly Pro Gly Leu Asp Ala Gly Ser Pro
1             5             10             15
Gly Val Pro Gly Trp Arg Pro Trp Arg Met Gly Trp Arg Pro Pro Thr
      20             25             30
Glu Gln Ala Gly Val Tyr Ser Ser Leu Phe Trp Glu Cys Ser Val Gly
      35             40             45
Gly Tyr Ala Leu Gly Val Trp Ser Leu Glu Pro His Trp Asp Arg Val
      50             55             60
Gln Cys Trp Pro Arg Leu Ser Arg Val Ala Gly Ile Phe Leu Arg Arg
65             70             75             80
Asn Gln Ser Cys Ser Glu Val Cys Cys Ser Ser Val Gly Leu Pro Trp
      85             90             95
Ala Ala Arg Ala Gly Gly Met Trp Glu Gly Ala Pro Asp Met His Leu

```


100 105 110
 Gly Ser Ser Ser Leu Gln Pro Thr Thr Gln Arg Ser
 115 120
 <210> 223
 <211> 331
 <212> DNA
 <213> Homo sapiens
 <400> 223
 tcattgaaatc tgtgggcagt gacccaggag ggtatgggca ggcccaacca ggttggtgtg
 60
 cccttgaagc cccacagacc tgccagggca gcagggcagt tgggagccgg agaacctgag
 120
 aaccaagcca ggctgcatgc aggaggctgg cacgtgaacg ctgcaggtgt tgccggcagc
 180
 cgtggtgcct ggcagatagt gtccgacccc cnaggacctt cttgctgggc agcccagtc
 240
 aaaagctggt cccgcttaag ccacccccac cgccttgccc acacctggca catgggtgaa
 300
 gcaagggcat ttcccggggc ttctgttcc c
 331

<210> 224
 <211> 103
 <212> PRT
 <213> Homo sapiens

<400> 224
 Met Pro Leu Leu His Pro Cys Ala Arg Cys Gly Gln Gly Gly Gly Gly
 1 5 10 15
 Gly Leu Ser Gly Asn Ser Phe Trp Thr Gly Leu Pro Ser Lys Lys Val
 20 25 30
 Leu Gly Gly Arg Thr Leu Ser Ala Arg His His Gly Cys Arg Gln His
 35 40 45
 Leu Gln Arg Ser Arg Ala Ser Leu Leu His Ala Ala Trp Leu Gly Ser
 50 55 60
 Gln Val Leu Arg Leu Pro Thr Ala Leu Leu Pro Trp Gln Val Cys Gly
 65 70 75 80
 Ala Ser Arg Ala His Gln Pro Gly Trp Ala Cys Pro Tyr Pro Pro Gly
 85 90 95
 Ser Leu Pro Thr Asp Phe Met
 100

<210> 225
 <211> 339
 <212> DNA
 <213> Homo sapiens

<400> 225
 tgatcacggg cgtgagccac cagcccagca tcccttgctt ttcattcgca cctccacctc
 60
 cagaatgacc ctcatccctt cctgcacaga cggtgacagc agtaactcct acaaacacca
 120

ccagactgat cttcaagagc agaggaactc ccaatcacga ttccaccccc gccgggctct
 180
 caaatcctcc agggctgect gctatggggg agggaggcac actttgcttg gctctcaagg
 240
 cctcagccag ccgggtccaa accaactccc agcctggcct caccatccca ccgccaacc
 300
 ttgctcaca ctggccccctc ttcttgggaac atgggcctn
 339

<210> 226

<211> 91

<212> PRT

<213> Homo sapiens

<400> 226

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Thr | Leu | Ile | Pro | Ser | Cys | Thr | Asp | Gly | Asp | Ser | Ser | Asn | Ser | Tyr |
| 1 | | | | 5 | | | | | 10 | | | | | 15 | |
| Lys | His | His | Gln | Thr | Asp | Leu | Gln | Glu | Gln | Arg | Asn | Ser | Gln | Ser | Arg |
| | | | 20 | | | | | 25 | | | | | 30 | | |
| Phe | His | Pro | Arg | Arg | Ala | Leu | Lys | Ser | Ser | Arg | Ala | Ala | Cys | Tyr | Gly |
| | | 35 | | | | | 40 | | | | | 45 | | | |
| Gly | Gly | Arg | His | Thr | Leu | Leu | Gly | Ser | Gln | Gly | Leu | Ser | Gln | Pro | Gly |
| | 50 | | | | | 55 | | | | | 60 | | | | |
| Pro | Asn | Gln | Leu | Pro | Ala | Trp | Pro | His | His | Pro | Thr | Ala | Lys | Pro | Leu |
| 65 | | | | 70 | | | | | 75 | | | | | 80 | |
| Leu | Thr | Leu | Ala | Pro | Leu | Pro | Gly | Thr | Trp | Ala | | | | | |
| | | | | 85 | | | | | 90 | | | | | | |

<210> 227

<211> 353

<212> DNA

<213> Homo sapiens

<400> 227

gtcgaccctc tcgattgtgg cgaactccat ggctgctgcg ggctgctgta ggctctcgag
 60
 tagctcgacg tcgggttcgc gagggctcgc agcgtggcca tgctgcttct tggatgggtc
 120
 gggcaactcc tcgggggatt cgagcagttc ttggcgcacc tgctctggcg tcatcccgga
 180
 ggccaggccg acaagtgctg cctcctgcca cccgctgagc gacgctgcca tgttgagtac
 240
 ggcgtcttca ctggtcaggg cgagcgcggt atcgaccagg ttggcgtcca ggccgagaga
 300
 cagcatgtct gctcagtcgc ggtgatgact ggagtggcgg tctcctgcac ggg
 353

<210> 228

<211> 102

<212> PRT

<213> Homo sapiens

<400> 228

Met Leu Ser Leu Gly Leu Asp Ala Asn Leu Val Asp Thr Ala Leu Ala

```

      1             5             10             15
Leu Thr Ser Glu Asp Ala Val Leu Asn Met Ala Ala Ser Leu Ser Gly
      20             25             30
Trp Gln Glu Ala Ala Leu Val Gly Leu Ala Ser Gly Met Thr Pro Glu
      35             40             45
Gln Val Arg Gln Glu Leu Leu Glu Ser Pro Glu Glu Leu Pro Glu Pro
      50             55             60
Ser Lys Lys Gln His Gly His Ala Ala Ser Pro Arg Glu Pro Asp Val
      65             70             75             80
Glu Leu Leu Glu Ser Leu Arg Arg Pro Ala Ala Met Glu Phe Ala
      85             90             95
Thr Ile Glu Gly Val Asp
      100

```

<210> 229
 <211> 743
 <212> DNA
 <213> Homo sapiens

<400> 229
 nnggctaggg acacggcctc ctctcaaca ggcagtgcct gtgcaggctc aggggcatca
 60
 tcaaagataa cacagggtcg gtcaggggct gctggctgct cctgccccag gactggctcc
 120
 aggatgggca aggtctgctc cctggtagcc agggggagag ggggaaggag caccagggag
 180
 tgggccagca ggtgtggcat cggccaggag gagatggagg ccagcagcag ccaagaccag
 240
 agtaaagtgt ctgccccagg ggtgtcaca gcccaggacc gggtagttag aaagccagcc
 300
 cagcttggca ctcagcggag ccaggaggca gatgttcagg actgggagtt cagaaagagg
 360
 gattcccagg gcacttactc cagccgggat gcagaactcc aggaccagga attcggaaag
 420
 agagattcac tgggtaccta cagtagtcca gatgtaagcc ttggggactg ggaatttggg
 480
 aagagagatt ctctgggtgc ttatgccagc caagatgcca acgagcaggg ccaagatttg
 540
 gggaagaggg accaccatgg taggtacagc agccaggatg ccgatgagca ggactgggag
 600
 tttcagaaga gagatgtgtc actcggcacc tatggcagcc gggctgcgga gccacaggaa
 660
 caggagtttg ggaagagcgc ttggataagg gactacagca gtggtggcag ctccaggacc
 720
 cttgacgccc aggacagaag ctt
 743

<210> 230
 <211> 247
 <212> PRT
 <213> Homo sapiens

<400> 230
 Xaa Ala Arg Asp Thr Ala Ser Ser Ser Thr Gly Ser Ala Cys Ala Gly

1 5 10 15
 Ser Gly Ala Ser Ser Lys Ile Thr Gln Gly Trp Ser Gly Ala Ala Gly
 20 25 30
 Cys Ser Cys Pro Arg Thr Gly Ser Arg Met Gly Lys Ala Ala Ser Leu
 35 40 45
 Val Ala Arg Gly Arg Gly Glu Gly Ser Thr Arg Glu Trp Ala Ser Arg
 50 55 60
 Cys Gly Ile Gly Gln Glu Glu Met Glu Ala Ser Ser Ser Gln Asp Gln
 65 70 75 80
 Ser Lys Val Ser Ala Pro Gly Val Leu Thr Ala Gln Asp Arg Val Val
 85 90 95
 Gly Lys Pro Ala Gln Leu Gly Thr Gln Arg Ser Gln Glu Ala Asp Val
 100 105 110
 Gln Asp Trp Glu Phe Arg Lys Arg Asp Ser Gln Gly Thr Tyr Ser Ser
 115 120 125
 Arg Asp Ala Glu Leu Gln Asp Gln Glu Phe Gly Lys Arg Asp Ser Leu
 130 135 140
 Gly Thr Tyr Ser Ser Arg Asp Val Ser Leu Gly Asp Trp Glu Phe Gly
 145 150 155 160
 Lys Arg Asp Ser Leu Gly Ala Tyr Ala Ser Gln Asp Ala Asn Glu Gln
 165 170 175
 Gly Gln Asp Leu Gly Lys Arg Asp His His Gly Arg Tyr Ser Ser Gln
 180 185 190
 Asp Ala Asp Glu Gln Asp Trp Glu Phe Gln Lys Arg Asp Val Ser Leu
 195 200 205
 Gly Thr Tyr Gly Ser Arg Ala Ala Glu Pro Gln Glu Gln Glu Phe Gly
 210 215 220
 Lys Ser Ala Trp Ile Arg Asp Tyr Ser Ser Gly Gly Ser Ser Arg Thr
 225 230 235 240
 Leu Asp Ala Gln Asp Arg Ser
 245

<210> 231

<211> 431

<212> DNA

<213> Homo sapiens

<400> 231

acgcgttggc caccgagagg ctggcgaggg tgtgcagcac ggcgcagtgt ggcaggggtcc
 60
 caggggtgcag cctgcgcagc agctcctcca tcaccttgct gatgaactgt cttcccacgg
 120
 ccaccaggac gccactcgcc gcctgtgtgcc agtcccagac caggtccttc gtcttggtca
 180
 tctcgctgga ggccaggagg atgatgggtgc tggctgtgtc cttgtccagc tcttgggcgc
 240
 gactgtctcag gacctctccc atggccctca ggaccgctgc tcggtatggg tgtgccagct
 300
 tgtcatgctg ccgcagatac tcctcgcagg cacggagcgt ctccaccctg ctggacgcc
 360
 tcaccgataa ggaccccctg gtgcaggagc aggtctgcag tgccctgtgc tccctcgggg
 420
 aggtgcggcc g
 431

<210> 232
 <211> 120
 <212> PRT
 <213> Homo sapiens

<400> 232
 Met Ala Ser Ser Arg Val Glu Thr Leu Arg Ala Cys Glu Glu Tyr Leu
 1 5 10 15
 Arg Gln His Asp Lys Leu Ala His Pro Tyr Arg Ala Ala Val Leu Arg
 20 25 30
 Ala Met Glu Arg Val Leu Ser Ser Arg Ala Ser Glu Leu Asp Lys Asp
 35 40 45
 Thr Ala Ser Thr Ile Ile Leu Leu Ala Ser Ser Glu Met Thr Lys Thr
 50 55 60
 Lys Asp Leu Val Trp Asp Trp Gln Gln Ala Ala Ser Gly Val Leu Val
 65 70 75 80
 Ala Val Gly Arg Gln Phe Ile Ser Lys Val Met Glu Glu Leu Leu Arg
 85 90 95
 Arg Leu His Pro Gly Thr Leu Pro His Cys Ala Val Leu His Thr Leu
 100 105 110
 Ala Ser Leu Ser Val Ala Asn Ala
 115 120

<210> 233
 <211> 606
 <212> DNA
 <213> Homo sapiens

<400> 233
 acgcgttcag ggatgccaga aatctaactg ggtaataaaa agctgggaga acattccaga
 60
 aagggtgggca cccttagcat tcccaaaaag caccagccct cctcatcctt ccagcttct
 120
 gtgctggaat gcacccccat cggaaaggct cgaaaactca ggacacatta ggatcacctg
 180
 gaaagcattt gtcaaaacgc atctccctgc gggtcagggt ccaagttaa atcaaacttc
 240
 aggtgatgct gactcagggt gctccagaaa cacctgggga agcagcactt tggaggctgc
 300
 ctctcacatc caccacacag caagtgggca gggagctagg taaatctcct tccagttga
 360
 gaaggggctc ggagcaggca cagagaagag atacccttag aatgcaagtt gttcagctgc
 420
 gaaagtccag cctgcaggct tcctgggcaa gctagtgggc tgaagtatgc cacagcaaca
 480
 ggcttctaga gccggctgcc cagctcctac tctgcctctg ccactcactg actgtgtggt
 540
 cttgagcagg tcacctgtct gacttggtga gagctgacag gcatcacctg ttagaggctt
 600
 acgcgt
 606

<210> 234

<211> 108
 <212> PRT
 <213> Homo sapiens

<400> 234

```

Met His Pro His Arg Lys Gly Ser Lys Thr Gln Asp Thr Leu Gly Ser
 1             5             10             15
Pro Gly Lys His Leu Ser Lys Arg Ile Ser Leu Arg Val Arg Val Gln
      20             25             30
Val Lys Ile Lys Leu Gln Val Met Leu Thr Gln Val Ala Pro Glu Thr
      35             40             45
Pro Gly Glu Ala Ala Leu Trp Arg Leu Pro Leu Thr Ser Thr Pro Gln
      50             55             60
Gln Val Gly Arg Glu Leu Gly Lys Ser Pro Ser Gln Leu Arg Arg Gly
65             70             75             80
Ser Glu Gln Ala Gln Arg Arg Asp Thr Leu Arg Met Gln Val Val Gln
      85             90             95
Leu Arg Lys Ser Ser Leu Gln Ala Ser Trp Ala Ser
      100             105

```

<210> 235
 <211> 328
 <212> DNA
 <213> Homo sapiens

<400> 235

```

cgaccgttga ctattctcta caaaccacaa agacaatgat tgatttaact gaatttagaa
60
atagcaaaca cttaaaacag cagcagtaca gagctgaaaa ccagattctt ttgaaagaga
120
ttgaaagtct agaggaagaa cgacttgatc tgaaaaaaaa aattcgccaa atggctcaag
180
aaagaggaaa aagaagggca acttcaggat taaccactgg ggacctgaac ctaactgaaa
240
acatttctca aggagataga ataagtgaag gaaaattgga tttattgagc ctcaaaaata
300
tgagtgaagc acaatcaaag aatgaatt
328

```

<210> 236
 <211> 97
 <212> PRT
 <213> Homo sapiens

<400> 236

```

Met Ile Asp Leu Thr Glu Phe Arg Asn Ser Lys His Leu Lys Gln Gln
 1             5             10             15
Gln Tyr Arg Ala Glu Asn Gln Ile Leu Leu Lys Glu Ile Glu Ser Leu
      20             25             30
Glu Glu Glu Arg Leu Asp Leu Lys Lys Lys Ile Arg Gln Met Ala Gln
      35             40             45
Glu Arg Gly Lys Arg Arg Ala Thr Ser Gly Leu Thr Thr Gly Asp Leu
      50             55             60
Asn Leu Thr Glu Asn Ile Ser Gln Gly Asp Arg Ile Ser Glu Arg Lys

```

65 70 75 80
 Leu Asp Leu Leu Ser Leu Lys Asn Met Ser Glu Ala Gln Ser Lys Asn
 85 90 95
 Glu

<210> 237

<211> 2059

<212> DNA

<213> Homo sapiens

<400> 237

ggccataagg gcacgacgca ttcctagccg atgcaccaac acgggcatga agcctgccga
 60
 gagcagaag ccggcgcca tagctacggc ccatacggtc atgtctgcca tggctccgtt
 120
 gatgtcagac tgcacatgaa atcggttacg gtacccagg atcatcgcta ccgagtacac
 180
 cccgaacagc acccgctggg cgccgatcag cgtgagggag tgccccacca gtggcacttt
 240
 tcttagatag cggaacccat ccaccacatc cccagtcacc gttctcatcg tccgggaacg
 300
 atccaccagt ggcggcccaa gctcccgacg tgaaaactgc agcccctagg cgaccgagac
 360
 tgcgaagagg gctgcggaga tgcagaaaat gatcgtgtcg gcgtggtgca caggaatatg
 420
 gcgtccggca atcatgcgca ctgctgcagc aacaaccgca ccgatcatga gccctagcgg
 480
 ccaatcgctg gcatgattga cgatgccgtc aggtagtcgc gcttgctgat ggtgtattcc
 540
 aaccagcga ccaaggcggg gagcaaaaac cggttcaggc tcatcgcgat gagcaacca
 600
 atgagcaagg ccaggtggga gggcttatcg cgcgcaccac ccagaccaa gatccccagc
 660
 ccgaccagg tgacggcacg cattcatctg cgtattgtcc cgactacacc gtgagggcgc
 720
 tctctgatct gcagctcatc aagggttacg gactgcagta cctcaatgca ctctggcta
 780
 cccgagccca gaacctgcca cagtcccctg agaacaccga cctgcagggt attccaggca
 840
 gccagaccag gtccttggt gagaagacca ccacagcggc agctttccca gtagcccttt
 900
 ccctcttttg cacagttgga acctccagtt gataaatgac tgtggactag cgcgcgtttt
 960
 ttgttttcag agcacacgta aggggtccagc cacagcaggc ccggcgcccc ggtggaaggc
 1020
 agccctgggc ggaaccagg cgtttaacgg ctactaggc agccccagat ctggggaagg
 1080
 agatgagcac gtggggagct ggagtgagct gagcagaagt tttgtgccc cctgccccca
 1140
 tcccctccag gccacgtttt agatggccct tgtagttgcg ggtcctgggt gtcctcagaa
 1200
 ctagacatca atgcctggat ccttcagccg gccctgccct cctttaggag acaggagtca
 1260

ccagggcaca gccctccagg cccgcctcag gaaggaatga aaggaatgcc atcatctcta
 1320
 gttcccaggg ccagccttc cctttctccc cggggcagg gacagtgcgg catattcaga
 1380
 ttcagacctc tttgggctga gccaccttgt gagtgcagtt actgcctttg tgtggccgtg
 1440
 acctctatTT gtttgctttt aatttgccaa cctatcgctg ctggcagcac tttttgagca
 1500
 agccgagagc acccattttg gctggggatt cagatcgatg gccttgtecca tgttgctcctt
 1560
 tctggcttcc ctgatgggtg catgtttcag cgcatgcgcc ccagccttcc ccatgtgccca
 1620
 aaccagaagc tccactgccc gtaggctgtc cctgtagccc tgctccctcc ctggaggctg
 1680
 ctcttctgat tctgagagct ggcctagtgg tgctgagggc ccctttctgc ttctctgccc
 1740
 acctgctgag ttgccactcg cagtgttgct agttcccgctg ttctgagaag aggtcatgcc
 1800
 tgggaggaag ggatcgatc gctgcatcga atctctcttc cgccgtgtgg cccccaggag
 1860
 agtagctgcc tgttgacact gctccacacc tccccacagc ctccctgcag gtgctgtgtg
 1920
 gccgtgatgt gcagagagca gtgagggagg gttcatgaac caggtggatc ctctttaaaa
 1980
 aaaaaaaaaag tttttgttat atctctaaaa tcccatagct aggaacagaa aaaaaggaaa
 2040
 agacttgaaa tgttctaga
 2059

<210> 238

<211> 129

<212> PRT

<213> Homo sapiens

<400> 238

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Ala | Glu | Gln | Lys | Phe | Cys | Ala | Arg | Leu | Pro | Pro | Ser | Pro | Pro | Gly | His |
| 1 | | | | 5 | | | | 10 | | | | | | 15 | |
| Val | Leu | Asp | Gly | Pro | Cys | Ser | Cys | Gly | Ser | Trp | Val | Ser | Ser | Glu | Leu |
| | | 20 | | | | | | 25 | | | | | | 30 | |
| Asp | Ile | Asn | Ala | Trp | Ile | Leu | Gln | Pro | Ala | Leu | Pro | Ser | Phe | Arg | Arg |
| | | 35 | | | | | 40 | | | | | 45 | | | |
| Gln | Glu | Ser | Pro | Gly | His | Ser | Pro | Pro | Gly | Pro | Pro | Gln | Glu | Gly | Met |
| | 50 | | | | | 55 | | | | | 60 | | | | |
| Lys | Gly | Met | Pro | Ser | Ser | Leu | Val | Pro | Arg | Ala | Gln | Pro | Ser | Pro | Ser |
| 65 | | | | 70 | | | | | 75 | | | | | 80 | |
| Pro | Pro | Gly | Gln | Gly | Gln | Cys | Gly | Ile | Phe | Arg | Phe | Arg | Pro | Leu | Trp |
| | | | 85 | | | | 90 | | | | | | | 95 | |
| Ala | Glu | Pro | Pro | Cys | Glu | Cys | Ser | Tyr | Cys | Leu | Cys | Val | Ala | Val | Thr |
| | | 100 | | | | | 105 | | | | | 110 | | | |
| Ser | Ile | Cys | Leu | Leu | Leu | Ile | Cys | Gln | Pro | Ile | Ala | Ala | Gly | Ser | Thr |
| | | 115 | | | | | 120 | | | | | 125 | | | |
| Phe | | | | | | | | | | | | | | | |

<210> 239
 <211> 388
 <212> DNA
 <213> Homo sapiens

<400> 239
 ntctagatca ctctgtagcg catgggttaa tgctgacaca atagaaaagt gcgaggacat
 60
 cctcgaatta atgagatggg ggactggatg agtcaagtgc tcgtcgttgc ggcggctgtc
 120
 ggtcagctgc cctcctcca cttctgcttc tcggcggttac cccataccgt attggccgcg
 180
 tgttcacctt tgaatgcagc catgtcgtcg tctccgtatc gaaatgatgt gccatcgaag
 240
 atgccgacct cagcatcggc atctgcagtg atgagtgcgt atcgcgccac acgaaacgcc
 300
 cagcgcaacc gtgtcctcgc acgatacgaa gtgcttgggt atctcagctc tggtagctat
 360
 ggtagctgtat ataaagcaaa ggaacttn
 388

<210> 240
 <211> 104
 <212> PRT
 <213> Homo sapiens

<400> 240
 Met Val Asp Trp Met Ser Gln Val Leu Val Val Ala Ala Ala Val Gly
 1 5 10 15
 Gln Leu Pro Leu Leu His Phe Cys Phe Ser Ala Leu Pro His Thr Val
 20 25 30
 Leu Ala Ala Cys Ser Pro Leu Asn Ala Ala Met Ser Ser Ser Pro Tyr
 35 40 45
 Arg Asn Asp Val Pro Ser Lys Met Pro Thr Ser Ala Ser Ala Ser Ala
 50 55 60
 Val Met Ser Ala Tyr Arg Ala Thr Arg Asn Ala Gln Arg Asn Arg Val
 65 70 75 80
 Leu Ala Arg Tyr Glu Val Leu Gly Tyr Leu Ser Ser Gly Thr Tyr Gly
 85 90 95
 Arg Val Tyr Lys Ala Lys Glu Leu
 100

<210> 241
 <211> 330
 <212> DNA
 <213> Homo sapiens

<400> 241
 ncggggggcc gagttgaaag ctgccggcac actggctgtg ctgcttgctt cacttctcgg
 60
 gatgctgctt ccagggcggg cctgggggaa acatcggcct tcccaggcac ccttagcccg
 120
 tcccatctgg gggcccttag cacagtcctt gggacccac atgctgcctt tcaggctgat
 180

gtgggcaaac tcggcagccc agcctactcc cgggccatgg gccaccatct cagcttcctt
 240
 ggggctaagc cgtgtgctct gaatcaaaag cagtagtggc atcggcggca ctggcgccat
 300
 gggaaacggg ttgacttgca caaccagcac
 330

<210> 242
 <211> 100
 <212> PRT
 <213> Homo sapiens

<400> 242
 Met Ala Pro Val Pro Pro Met Pro Leu Leu Leu Leu Ile Gln Ser Thr
 1 5 10 15
 Arg Leu Ser Pro Arg Glu Ala Glu Met Val Ala His Gly Pro Gly Val
 20 25 30
 Gly Trp Ala Ala Glu Phe Ala His Ile Ser Leu Lys Gly Ser Met Trp
 35 40 45
 Gly Pro Arg Asp Cys Ala Lys Gly Pro Gln Met Gly Arg Ala Lys Gly
 50 55 60
 Ala Trp Glu Gly Arg Cys Phe Pro Gln Ala Arg Pro Gly Ser Ser Ile
 65 70 75 80
 Pro Arg Ser Glu Ala Ser Ser Thr Ala Ser Val Pro Ala Ala Phe Asn
 85 90 95
 Ser Ala Pro Arg
 100

<210> 243
 <211> 330
 <212> DNA
 <213> Homo sapiens

<400> 243
 nnaccttctc tccgcgttat taccaaagat gctatgcacg taactgcgga ggaaattctt
 60
 cacacaggcc accccgcccc cactgcgctc gtcgctaate ttccctataa cggtgcggta
 120
 cccgtactgc tacacatgct agatattctc ccctccttgc ggactacagt ggtgatggtg
 180
 caggcagaag tagccgatcg attggctgcc acaccaggca gccgcattta cgggtgtcccc
 240
 agcgtcaaag tcaactttta cgggactgtc tcgctgctgg gagcaattgg acgcaatgtc
 300
 ttctggccgg ctcccaatgt tgattctggn
 330

<210> 244
 <211> 110
 <212> PRT
 <213> Homo sapiens

<400> 244
 Xaa Pro Ser Leu Arg Val Ile Thr Lys Asp Ala Met His Val Thr Ala

```

      1           5           10           15
Glu Glu Ile Leu His Thr Gly His Pro Ala Pro Thr Ala Leu Val Ala
      20           25           30
Asn Leu Pro Tyr Asn Val Ala Val Pro Val Leu Leu His Met Leu Asp
      35           40           45
Ile Leu Pro Ser Leu Arg Thr Thr Val Val Met Val Gln Ala Glu Val
      50           55           60
Ala Asp Arg Leu Ala Ala Thr Pro Gly Ser Arg Ile Tyr Gly Val Pro
65      70           75           80
Ser Val Lys Val Asn Phe Tyr Gly Thr Val Ser Arg Ala Gly Ala Ile
      85           90           95
Gly Arg Asn Val Phe Trp Pro Ala Pro Asn Val Asp Ser Gly
      100           105           110

```

<210> 245
 <211> 355
 <212> DNA
 <213> Homo sapiens

```

<400> 245
tctagatcct gaatcaccca cctcctagtt tcggattcac ctccgccggc gtcacctgaa
60
aacaatgtcg agcccgaatg gatgatggta gccacaccca tctcggaaag gtggaatgca
120
gcgtgttgca gaaacagaag ttgaccgtcg gaggtaggcg gcattcgctt cggatcgaag
180
cgtcccgagg catccatctc gagttgacga cgaaaatctt tccagtccac gccgtagggg
240
ganttgcaa ccacagcatc gaatttgtcc agaaggaagt ggtcgttggt gagggatttg
300
ccccattcaa tacgcgcatc ttcccgaag cgcgcctcta ttgcggccaa cgcgt
355

```

<210> 246
 <211> 101
 <212> PRT
 <213> Homo sapiens

```

<400> 246
Met Arg Val Leu Asn Gly Ala Ile Pro Ser Pro Thr Thr Thr Ser Phe
1           5           10           15
Trp Thr Asn Ser Met Leu Trp Leu Pro Xaa Pro Pro Thr Ala Trp Thr
      20           25           30
Gly Lys Ile Phe Val Val Asn Ser Arg Trp Met Pro Arg Asp Ala Ser
      35           40           45
Ile Arg Ser Glu Cys Arg Leu Pro Pro Thr Val Asn Phe Cys Phe Cys
      50           55           60
Asn Thr Leu His Ser Thr Phe Pro Arg Trp Val Trp Leu Pro Ser Ser
65      70           75           80
Ile Arg Ala Arg His Cys Phe Gln Val Thr Pro Ala Glu Val Asn Pro
      85           90           95
Lys Leu Gly Gly Gly
      100

```

<210> 247
 <211> 333
 <212> DNA
 <213> Homo sapiens

<400> 247
 atggccgcga atgggcaccg tgatcatggc gtctctcccc gctacgacca gtacaaggac
 60
 gcctgggaca ccagcgtcgt gtccgagatc aagatgggag acaggtacga gacggtcagg
 120
 ttcttccact gctacaagcg cggagtggac cgcgtgttcg ttgaccaccc actgttctctg
 180
 gagagggttt ggggaaagac cgaggagaag atctacgggc ctgacgctgg aacggactac
 240
 agggacaacc agctgcggtt cagcctgcta tgccaggcag cacttgaagc tccaaggatc
 300
 ctgagcctca acaacaaccc atacttctcc gga
 333

<210> 248
 <211> 111
 <212> PRT
 <213> Homo sapiens

<400> 248
 Met Ala Ala Asn Gly His Arg Val Met Val Val Ser Pro Arg Tyr Asp
 1 5 10 15
 Gln Tyr Lys Asp Ala Trp Asp Thr Ser Val Val Ser Glu Ile Lys Met
 20 25 30
 Gly Asp Arg Tyr Glu Thr Val Arg Phe Phe His Cys Tyr Lys Arg Gly
 35 40 45
 Val Asp Arg Val Phe Val Asp His Pro Leu Phe Leu Glu Arg Val Trp
 50 55 60
 Gly Lys Thr Glu Glu Lys Ile Tyr Gly Pro Asp Ala Gly Thr Asp Tyr
 65 70 75 80
 Arg Asp Asn Gln Leu Arg Phe Ser Leu Leu Cys Gln Ala Ala Leu Glu
 85 90 95
 Ala Pro Arg Ile Leu Ser Leu Asn Asn Asn Pro Tyr Phe Ser Gly
 100 105 110

<210> 249
 <211> 5503
 <212> DNA
 <213> Homo sapiens

<400> 249
 atgaccagg ggattttggc cttggtcacg tccactggct gtgcatctgc caatgccctg
 60
 cagtcctca cggatgccat gcacatccca cacctctttg tccagcgcaa cccgggaggg
 120
 tcgccacgca ccgcatgcc cctgaacccc agccccgatg gtgaggccta cacactggct
 180
 tcgagaccac ccgtccgcct caatgatgtc atgtcaggc tggtagcgga gctgcgctgg
 240

cagaagttcg tcatgttcta cgacagcgag tatgatatcc gtgggcttca aagctttctg
300
gaccaggcct cgcggtggg ccttgacgtc tctttacaaa aggtggacaa gaacattagc
360
cacgtattca ccagcctgtt caccacgatg aagacagagg agctgaatcg ctaccgggac
420
acgttcgcc ggcctatcct gctgctcagc ccacagggag cccactcctt catcaacgag
480
gccgtggaga ccaacctggc ttccaaggac agccactggg tctttgtgaa tgaggaaatc
540
agtgaccggg agatcctgga tctgggtccat agtgcccttg gaaggatgac cgtgggtccg
600
caaatctttc cgtctgcaaa ggacaatcag aaatgcacga ggaacaacca ccgcatctcc
660
tccctgtctc gcgaccccca ggaaggctac ctccagatgc tgcagatctc caacctctat
720
ctgtatgaca gtgttctgat gctggccaac gcctttcaca ggaagctgga ggaccggaag
780
tggcatagca tggcgagcct caactgcata cggaaatcca ctaagccatg gaatggtggg
840
aggtccatgc tggataccat caaaaagggc cacatcactg gcctcactgg ggtgatggag
900
tttcgggagg acagttcgaa tccctatgtc cagtttgaaa tccctggcac tacctatagt
960
gagacttttg gcaaagacat gcgcaagttg gcgacatggg actcagagaa gggcttgaat
1020
ggcagcttgc aagagaggcc catgggcagc cgcctccaag gattgactct taaagtgggt
1080
actgtcttgg aagagccttt cgtgatggtg gctgagaaca tcctaggaca gccaagcgc
1140
tacaagggt tctccataga tgcctggat gcaactggcca aggtcttggg ctttaaata
1200
gagatttacc aagccctga tggcaggtac ggtcaccagc tccataacac ctcttggaa
1260
gggatgatcg gggagctcat cagcaagaga gcagacttgg ccatctctgc catcaccatc
1320
accccagaga gggagagcgt tgtggacttc agcaagcggg acatggacta ttcagtggg
1380
attctaatta agaagcccga ggagaaaatc agcatcttct ccctcttgc tccatttgat
1440
ttcgtgtgt gggcctgcat tgcagcagcc atccctgtgg ttggtgtgct gatatttgtg
1500
ttgaacagga tacaggctgt gagggctcag agtgctgccc agcccaggcc gtcagcttct
1560
gccactctgc acagcgccat ctggattgtc tatggagcct tcgtacagca aggtggcgaa
1620
tcttcctga actccatggc catgcgcac gtgatgggca gctgggtggc cttcacgctc
1680
attgtgtgct cctctacac agccaacctt gctgccttcc tcacagtgtc caggatggac
1740
aaccataa ggactttcca ggacctgtcc aaacaagtgg aaatgtctta tggcactgtc
1800
cgggattctg ctgtatatga gtacttccga gccaaaggga ccaacccctt ggagcaggac
1860

agcacgtttg ctgaactctg gcggaccatc agcaagaacg gaggggctga caactgcgtg
1920
tccagtcctt cagaaggcat caggaaggca aagaagggga actacgcctt cctgtgggat
1980
gtggccgtgg tggaatacgc agccctgacg gatgacgact gctcggtgac tgtcatcggc
2040
aacagcatca gcagcaaggg ttacgggatt gccctgcagc atggcagccc ctacagggac
2100
ctcttctccc agaggatcct ggagctgcag gacacagggg acctggatgt gctgaagcag
2160
aagtgggtggc cgcacatggg ccgctgtgac ctcaccagcc atgccagcgc ccaggccgac
2220
ggcaaatccc tcaagctgca cagcttcgcc ggggtcttct gcatcctggc cattggcctg
2280
ctcctggcct gcctgggtggc tgccctggag ttgtgggtgga acagcaaccg gtgccaccag
2340
gagaccccca aggaggacaa agaagtgaac ttggagcagg tccaccggcg catgaacagc
2400
ctcatggatg aagacattgc tcacaagcag atttccccag cgtcgattga gctctcggcc
2460
ctggagatgg ggggcctggc tcccaccag acctggagc cgacacggga gtaccagaac
2520
accagctct cggtcagcac ctttctgcc gagcagagca gccatggcac cagccggaca
2580
ctctcatcag ggcccagcag caacctgccg ctgccgctga gcagctcggc gaccatgcc
2640
tccatgcagt gcaaacacag gtcaccaac ggggggctgt tccggcagag cccgggtgaag
2700
accccatcc ccatgtcctt ccagcccgtg cctggaggcg tccttcaga ggctctggac
2760
acctccacg ggacctccat ctgactgcgc cgctgcctt cctgcccacc ctcccacca
2820
cccaccagc agagcttttt aatacaagaa aacaacaaca caaacacac aactcgcac
2880
acacacacat acacagagac tctttcattt ttcttgata tatgtgtaa taatgacaga
2940
atggagtggg gtaaaagtgt attttgaata ttccaattt tcgaagtcag taaaaaaca
3000
caaaaactgt atgaatgact ttgtaaattt tgttctatat gaataaaaag gcaaattact
3060
tgtgatcatt ctgaagtgcc aaaggagccc cccattcct gggccttct gagggcagga
3120
ggggcgacca gataaggagc ccctctctgc tgggggagaa ggaagacga ggaacccac
3180
atgccactcg ctgccttgtc ccacagcttg ctgcccatt tctttgctcc tggcacctcg
3240
tcccttttag tccctcagct tgataaagag tgagtttga gcccgattg ggctggcca
3300
ctgggttgc gtgctgtagg gtgatcggct gttctggta gcctggggct gaggaggtgc
3360
cctggactca gggctatcct gtccctgtctt ggatcttgcg ggacgagta gtcaccgctg
3420
tgtgtgttc agtgtgcctc tgccccatgg gcccgagaga agctgacaat taccatagt
3480

gaggtaattg agggccttga tactgagctg ttctcatatc atactgttg ccttgttttt
3540
cattttcttt tcaaaattag aagaatcaaa aagtgataga atattggggg aaggagggca
3600
gactccatcg ttctctcagg gagactgggc ggaggtgggg ttggtctgga agaaggtcca
3660
tgggggcagg gagttgagag gtgggggttag ttgcatggac caggtgaggt ggtggaacaa
3720
aaggccagggt agaggaagaa tattcccttg ggtttggacc catggtccca ggtgagagaa
3780
ggaagtgagg ccatagatgc agggagtaga agctttgttg tgtcaggcaa acggactctg
3840
ttagtaagag ttgggggagg tgcccagggt ggtggaccag tactccagggt gatggggcat
3900
gctgagaaat agagaaaaga gaccatgttt atttgggtag gaggaagcct tgccttgccg
3960
cttaggtgag aagcataagt gtaactccca tccaccaggg aagttgcttg tagcccaaaa
4020
agtaaaggcc tatctctggg tcataaatcc tgcaggcagt ccaacaaaca gggctggctc
4080
cagcacaac tctcccttcc acctttacga cctctccag accagacctg gagtctctt
4140
caagcagcaa tccaaccag agcaggggccc ctcccactc aggcactctga taacctctga
4200
gatttccagg cctatccctg tgcaggtaga gctgcaggga agcccagttt tctaagccta
4260
ggaccaccta acagcgcccc ctcttccagt tgccctttct gaggcaacta aactacagaa
4320
tcagggaag aaccattagg agtggaaactg ggggaatctg agttgtgtca cttcagttcc
4380
tctcctaaag acaaagggtta gtctgatctc cagaccgctc agaatggaat acacagccca
4440
catctgtcac tgaaggaggt ggagctccca cagccagcag taatcaggga gctgagagcc
4500
tggttctgtg atcatggaga aatacaaatg cctattgatt gcttcttcta tagccttgta
4560
gagtttctag agagatgtat ttatgagggt gataactagc ccaggattga tttctttcct
4620
aaagtcccta gtgacatgat tgagcagtaa agaacggcca aatcacacag tcagctaaaa
4680
gcactgtggg gaagagagtg ttataatta tgttatttat tgctggatgc tgagaatggt
4740
ctgatattcg tgctacctag gcaatccatt gacatttctc caatcagagc atgtggacct
4800
tgagaccagg catgctcaga gaagcctagg tgggctacca tgaccccgag gaagagcagg
4860
ctttgttttc catcagcacg ttgggggccc tgccctgaat ggtcaatttt tcacatatat
4920
atctctctat ttttttaate aaactctggt ctcaactgct tatctcacac caactctgtt
4980
tcctcattgc cccctgagat ggcctgtctt ctgggggtata gcttggatgt cttcttggat
5040
ggttctgctt agaatgagtg tcaaggagga aagagaggga gatggaggat gtgtttgtgc
5100

gcctgtgtgt gtgtgtttgt gtgtgtgtgt gtgtgtgaga gagagagaga gagagagaga
 5160
 gaccagcatc ttcaagagaa gtattctgct tatacaaaat ccttaacacc tcatgggtgtt
 5220
 attcttcacc atgtttatat atatatatat atattttttt ttttttttag aattttctac
 5280
 ccttggcatg aggggaaatg attgatattc aagcaagttc tctaggaaaa aaaaaaaact
 5340
 tcccaactca gatttctgtg tcagctcaga atgtatcttt ttttcatgct ttgctctttg
 5400
 gatttataac tctgtttaga ctattccata catttttaggt atattttgtg ccttcagaca
 5460
 ctgcaaataa taatcagcat ttggattaaa gttgtttaat aat
 5503

<210> 250

<211> 927

<212> PRT

<213> Homo sapiens

<400> 250

| | | | | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Thr | Gln | Gly | Ile | Leu | Ala | Leu | Val | Thr | Ser | Thr | Gly | Cys | Ala | Ser | 1 | 5 | 10 | 15 |
| Ala | Asn | Ala | Leu | Gln | Ser | Leu | Thr | Asp | Ala | Met | His | Ile | Pro | His | Leu | 20 | 25 | 30 | |
| Phe | Val | Gln | Arg | Asn | Pro | Gly | Gly | Ser | Pro | Arg | Thr | Ala | Cys | His | Leu | 35 | 40 | 45 | |
| Asn | Pro | Ser | Pro | Asp | Gly | Glu | Ala | Tyr | Thr | Leu | Ala | Ser | Arg | Pro | Pro | 50 | 55 | 60 | |
| Val | Arg | Leu | Asn | Asp | Val | Met | Leu | Arg | Leu | Val | Thr | Glu | Leu | Arg | Trp | 65 | 70 | 75 | 80 |
| Gln | Lys | Phe | Val | Met | Phe | Tyr | Asp | Ser | Glu | Tyr | Asp | Ile | Arg | Gly | Leu | 85 | 90 | 95 | |
| Gln | Ser | Phe | Leu | Asp | Gln | Ala | Ser | Arg | Leu | Gly | Leu | Asp | Val | Ser | Leu | 100 | 105 | 110 | |
| Gln | Lys | Val | Asp | Lys | Asn | Ile | Ser | His | Val | Phe | Thr | Ser | Leu | Phe | Thr | 115 | 120 | 125 | |
| Thr | Met | Lys | Thr | Glu | Glu | Leu | Asn | Arg | Tyr | Arg | Asp | Thr | Leu | Arg | Arg | 130 | 135 | 140 | |
| Ala | Ile | Leu | Leu | Leu | Ser | Pro | Gln | Gly | Ala | His | Ser | Phe | Ile | Asn | Glu | 145 | 150 | 155 | 160 |
| Ala | Val | Glu | Thr | Asn | Leu | Ala | Ser | Lys | Asp | Ser | His | Trp | Val | Phe | Val | 165 | 170 | 175 | |
| Asn | Glu | Glu | Ile | Ser | Asp | Pro | Glu | Ile | Leu | Asp | Leu | Val | His | Ser | Ala | 180 | 185 | 190 | |
| Leu | Gly | Arg | Met | Thr | Val | Val | Arg | Gln | Ile | Phe | Pro | Ser | Ala | Lys | Asp | 195 | 200 | 205 | |
| Asn | Gln | Lys | Cys | Thr | Arg | Asn | Asn | His | Arg | Ile | Ser | Ser | Leu | Leu | Cys | 210 | 215 | 220 | |
| Asp | Pro | Gln | Glu | Gly | Tyr | Leu | Gln | Met | Leu | Gln | Ile | Ser | Asn | Leu | Tyr | 225 | 230 | 235 | 240 |
| Leu | Tyr | Asp | Ser | Val | Leu | Met | Leu | Ala | Asn | Ala | Phe | His | Arg | Lys | Leu | 245 | 250 | 255 | |
| Glu | Asp | Arg | Lys | Trp | His | Ser | Met | Ala | Ser | Leu | Asn | Cys | Ile | Arg | Lys | | | | |

563

690 695 700
 Arg Ile Leu Glu Leu Gln Asp Thr Gly Asp Leu Asp Val Leu Lys Gln
 705 710 715 720
 Lys Trp Trp Pro His Met Gly Arg Cys Asp Leu Thr Ser His Ala Ser
 725 730 735
 Ala Gln Ala Asp Gly Lys Ser Leu Lys Leu His Ser Phe Ala Gly Val
 740 745 750
 Phe Cys Ile Leu Ala Ile Gly Leu Leu Leu Ala Cys Leu Val Ala Ala
 755 760 765
 Leu Glu Leu Trp Trp Asn Ser Asn Arg Cys His Gln Glu Thr Pro Lys
 770 775 780
 Glu Asp Lys Glu Val Asn Leu Glu Gln Val His Arg Arg Met Asn Ser
 785 790 795 800
 Leu Met Asp Glu Asp Ile Ala His Lys Gln Ile Ser Pro Ala Ser Ile
 805 810 815
 Glu Leu Ser Ala Leu Glu Met Gly Gly Leu Ala Pro Thr Gln Thr Leu
 820 825 830
 Glu Pro Thr Arg Glu Tyr Gln Asn Thr Gln Leu Ser Val Ser Thr Phe
 835 840 845
 Leu Pro Glu Gln Ser Ser His Gly Thr Ser Arg Thr Leu Ser Ser Gly
 850 855 860
 Pro Ser Ser Asn Leu Pro Leu Pro Leu Ser Ser Ser Ala Thr Met Pro
 865 870 875 880
 Ser Met Gln Cys Lys His Arg Ser Pro Asn Gly Gly Leu Phe Arg Gln
 885 890 895
 Ser Pro Val Lys Thr Pro Ile Pro Met Ser Phe Gln Pro Val Pro Gly
 900 905 910
 Gly Val Leu Pro Glu Ala Leu Asp Thr Ser His Gly Thr Ser Ile
 915 920 925

<210> 251

<211> 291

<212> DNA

<213> Homo sapiens

<400> 251

nngatcagcc gcgggggtccg cgccctcgat tcggcggtgg agaccgagag tctgcgtgag
 60
 gacgtcaacg cgctcgaacg gctgcggttg gccgtgcgcg ccagcgtggt catcctcatc
 120
 gagtaccacc attcgggtgac cctgctgctg cgggtgcgcg ggaactcacc tctggaacga
 180
 gaggccctcg aggcccgccg ccgtatcgat gcgaaggttc ccgctctcgt cgagagcgcc
 240
 atcgccgagg gtggtctgcg ctcggatttc actcccgggc tcatcacgcg t
 291

<210> 252

<211> 97

<212> PRT

<213> Homo sapiens

<400> 252

Xaa Ile Ser Arg Gly Val Arg Ala Leu Asp Ser Ala Val Glu Thr Glu

```

      1           5           10           15
Ser Leu Arg Glu Asp Val Asn Ala Leu Glu Arg Leu Arg Leu Ala Val
      20           25           30
Arg Ala Ser Val Val Ile Leu Ile Glu Tyr His His Ser Val Thr Leu
      35           40           45
Leu Leu Arg Val Arg Gly Asn Ser Pro Leu Glu Arg Glu Ala Leu Glu
      50           55           60
Ala Arg Arg Arg Ile Asp Ala Lys Val Pro Ala Leu Val Glu Ser Ala
      65           70           75           80
Ile Ala Glu Gly Gly Leu Arg Ser Asp Phe Thr Pro Gly Leu Ile Thr
      85           90           95
Arg

```

<210> 253
 <211> 327
 <212> DNA
 <213> Homo sapiens

```

<400> 253
gtgcacggat gggagcgcgc gcgcgcgtgc tggcgccttc acagcccggc gagcggcgtg
60
cgctcacggc cctgtaccga ccgatctcgc aaccttccgc agaccgatcc accaaccgcg
120
cccacatgtc ggcagtgatg gcgggcacct tgcgggagaa ggccgggaag gtcgagcgag
180
ccaatgaccg tcgcacggtc ggcacgctcc acgagcggga cgagaagctc gcggcaggac
240
gctcactcgt cgcggtgtcc tccgcggtct ccatcacctg ccctgcgaca tggaacgccc
300
acgacttcgg acggcgactc gacgcgt
327

```

<210> 254
 <211> 106
 <212> PRT
 <213> Homo sapiens

```

<400> 254
Met Gly Ala Leu Ala Arg Val Leu Val Pro Ser Gln Pro Gly Glu Arg
1           5           10           15
Arg Ala Leu Thr Val Leu Tyr Arg Pro Ile Ser Gln Pro Ser Ala Asp
      20           25           30
Arg Ser Thr Asn Arg Ala His Met Ser Ala Val Met Ala Gly Thr Leu
      35           40           45
Arg Glu Lys Ala Gly Lys Val Glu Arg Ala Asn Asp Arg Arg Thr Val
      50           55           60
Gly Thr Leu His Glu Arg Asp Glu Lys Leu Ala Gly Arg Ser Leu
      65           70           75           80
Val Ala Val Ser Ser Ala Val Ser Ile Thr Val Pro Ala Thr Trp Asn
      85           90           95
Ala His Asp Phe Gly Arg Arg Leu Asp Ala
      100           105

```

<210> 255
 <211> 372
 <212> DNA
 <213> Homo sapiens

<400> 255
 ctagaaatgg ctggctacga atacatggaa gctgaaaata gccacaagc ccacgaaatt
 60
 atcgtggacc atagacctga cttaatctta tgtgattgga tgatgccagg agggagtggc
 120
 atcgagctaa ctcgtcgctt aaagaaagac agcacgacag cagaaatccc tgttatttta
 180
 ctaacggcca aaagtgaaga agacaataaa attcaaggct tagaagtcgg tgcagatgac
 240
 tacatcacta aacctttctc tcctcgtgaa ctagtagcac gcctcaaggc ggtattacgc
 300
 cgagcgactc cacaaggtat tgatgatcct attgaaattg atggtttaac gcttgatccc
 360
 attagccaac gc
 372

<210> 256
 <211> 124
 <212> PRT
 <213> Homo sapiens

<400> 256
 Leu Glu Met Ala Gly Tyr Glu Tyr Met Glu Ala Glu Asn Ser Gln Gln
 1 5 10 15
 Ala His Glu Ile Val Asp His Arg Pro Asp Leu Ile Leu Cys Asp
 20 25 30
 Trp Met Met Pro Gly Gly Ser Gly Ile Glu Leu Thr Arg Arg Leu Lys
 35 40 45
 Lys Asp Ser Thr Thr Ala Glu Ile Pro Val Ile Leu Leu Thr Ala Lys
 50 55 60
 Ser Glu Glu Asp Asn Lys Ile Gln Gly Leu Glu Val Gly Ala Asp Asp
 65 70 75 80
 Tyr Ile Thr Lys Pro Phe Ser Pro Arg Glu Leu Val Ala Arg Leu Lys
 85 90 95
 Ala Val Leu Arg Arg Ala Thr Pro Gln Gly Ile Asp Asp Pro Ile Glu
 100 105 110
 Ile Asp Gly Leu Thr Leu Asp Pro Ile Ser Gln Arg
 115 120

<210> 257
 <211> 639
 <212> DNA
 <213> Homo sapiens

<400> 257
 nnacgcgtag cggtcgaggt tgcggacacc atgccgaac ccggcctgct cgccatcgag
 60
 gcacccatgg gacacggcaa gaccgaggcc gccctcatgt gcgcacaggt gctcgccgaa
 120

cggttcgggc tcggcggcat cttcttcggt ctaccgacga tggccacgtc caatcccatg
 180
 ttcggtcgag ttcgggaatg gctggacgct gtgccagcca aggaccgctc aagcatttcc
 240
 ctgggtcact cgaaagctgg actcaacgag gagtaccagc agctcatgcc gtggaacgcc
 300
 accatggccg tctacgacga aggtgcccgc acgcagcgtg aagcttcggc gatcgtccat
 360
 gagtggttct tgggcccga ggcgcgcgatc ctggccgacc acgtcgtcgg gaccatcgac
 420
 caggcactgt tcaccggtct caaagccaag catgtggtgt tacgccacct cggctctggcg
 480
 agcaaggctg tcatcattga tgaggtccac gccgcgcgacg tctatatgcg cgaatacctc
 540
 aaggctcgtc tcgaatggct cggcgccctac cgcacgccag tcatcctcat gtccgcgacg
 600
 ctgccaccgg cccaacgtca tgaactcgcg ctagcgtac
 639

<210> 258

<211> 213

<212> PRT

<213> Homo sapiens

<400> 258

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Xaa | Arg | Val | Ala | Val | Glu | Val | Ala | Asp | Thr | Met | Pro | Glu | Pro | Gly | Leu |
| 1 | | | 5 | | | | | | 10 | | | | | 15 | |
| Leu | Ala | Ile | Glu | Ala | Pro | Met | Gly | His | Gly | Lys | Thr | Glu | Ala | Ala | Leu |
| | | 20 | | | | | | 25 | | | | | 30 | | |
| Met | Cys | Ala | Gln | Val | Leu | Ala | Glu | Arg | Phe | Gly | Leu | Gly | Gly | Ile | Phe |
| | | 35 | | | | | 40 | | | | | 45 | | | |
| Phe | Gly | Leu | Pro | Thr | Met | Ala | Thr | Ser | Asn | Pro | Met | Phe | Gly | Arg | Val |
| | 50 | | | | | 55 | | | | | 60 | | | | |
| Arg | Glu | Trp | Leu | Asp | Ala | Val | Pro | Ala | Lys | Asp | Pro | Ser | Ser | Ile | Ser |
| 65 | | | | 70 | | | | | 75 | | | | | 80 | |
| Leu | Ala | His | Ser | Lys | Ala | Gly | Leu | Asn | Glu | Glu | Tyr | Gln | Gln | Leu | Met |
| | | | 85 | | | | | 90 | | | | | | 95 | |
| Pro | Trp | Asn | Ala | Thr | Met | Ala | Val | Tyr | Asp | Glu | Gly | Ala | Gly | Thr | Gln |
| | | 100 | | | | | | 105 | | | | | 110 | | |
| Arg | Glu | Ala | Ser | Ala | Ile | Val | His | Glu | Trp | Phe | Leu | Gly | Arg | Lys | Arg |
| | | 115 | | | | | 120 | | | | | 125 | | | |
| Ala | Ile | Leu | Ala | Asp | His | Val | Val | Gly | Thr | Ile | Asp | Gln | Ala | Leu | Phe |
| | 130 | | | | | 135 | | | | | 140 | | | | |
| Thr | Gly | Leu | Lys | Ala | Lys | His | Val | Val | Leu | Arg | His | Leu | Gly | Leu | Ala |
| 145 | | | | 150 | | | | | 155 | | | | | 160 | |
| Ser | Lys | Val | Val | Ile | Ile | Asp | Glu | Val | His | Ala | Ala | Asp | Val | Tyr | Met |
| | | | 165 | | | | | 170 | | | | | | 175 | |
| Arg | Glu | Tyr | Leu | Lys | Val | Val | Leu | Glu | Trp | Leu | Gly | Ala | Tyr | Arg | Thr |
| | | 180 | | | | | 185 | | | | | 190 | | | |
| Pro | Val | Ile | Leu | Met | Ser | Ala | Thr | Leu | Pro | Pro | Ala | Gln | Arg | His | Glu |
| | | 195 | | | | | 200 | | | | | 205 | | | |
| Leu | Ala | Leu | Ala | Tyr | | | | | | | | | | | |
| | | 210 | | | | | | | | | | | | | |

<210> 259
 <211> 252
 <212> DNA
 <213> Homo sapiens

<400> 259
 acgcgtgcac tgttgttatg catggtaacg tacacgtgtg cactgtgtgt ggtgtgcatg
 60
 ncatgggtgtg tgcacgtgtg cnaactgtgta tgcattggtaa tgtgcacgtg tgcancgtgtg
 120
 tgtnggtgtg tatgcatgng tgtgtgcacg tgtgcactgn agtgtgggggt gtatgcatgg
 180
 tgtgtgcaca tgagcactgt gtggtgtgta tgcattggtgn ggtgcacgtg tgcactgtgt
 240
 atgcaatggg gt
 252

<210> 260
 <211> 84
 <212> PRT
 <213> Homo sapiens

<400> 260
 Thr Arg Ala Leu Cys Val Cys Met Val Thr Tyr Thr Cys Ala Leu Cys
 1 5 10 15
 Val Val Cys Met Xaa Trp Cys Val His Val Cys Xaa Cys Val Cys Met
 20 25 30
 Val Met Cys Thr Cys Ala Xaa Val Cys Xaa Cys Val Cys Met Xaa Val
 35 40 45
 Cys Thr Cys Ala Leu Xaa Cys Gly Val Tyr Ala Trp Cys Val His Met
 50 55 60
 Ser Thr Val Trp Cys Val Cys Met Val Xaa Cys Thr Cys Ala Leu Cys
 65 70 75 80
 Met Gln Trp Cys

<210> 261
 <211> 1202
 <212> DNA
 <213> Homo sapiens

<400> 261
 gctagcccgg tcgcgttcgt cgtcgatttg ctggcggcag tcccctcgat cgtcttcggg
 60
 ctgtggggcg gcacgtctct cggatcgtcg ggaatcatca acggttacgc gggggcctta
 120
 ttcaaagcgc tcggctggat tccgatcttt tccgaagatc cgtcgtggtc ctccggtact
 180
 ggcacggctc accttgccag tctcgtcctg gccatcatga tccctgccaat tatcactgct
 240
 gttagccgcg acgtcatgcc ccgaacgccc catgatcaag tcgaggccgc gctcgccttc
 300
 ggatcgacgc gctgggaggt catcaagctt gcagtgttcc cccactcgcg gtccggcatc
 360

atttccggat ccattgttggg tctaggacgc gccctcggcg agaccctggc tgtcacccctc
 420
 atcctgcaga cgatgagccc catggcgctc aaacagaacc tcaacctgtc gatcttcgtc
 480
 ggtggtgaga cattcgcgctc gaagattgcc ggtaacttct ccgaggccat tagcgatccc
 540
 acctcgctgg gtgccctcgt ggcgtcggcc ctggccctgt tcgtcattac cttcgtggtc
 600
 aacgcgactg cccggttgat tgcggcgaag ggggttaagc gatgagcgcc accaccctg
 660
 accacatcac ccaccatggc gacaacacgc ccggacagct agatctctcc cgccgtctg
 720
 gtaaacggac tatcaagagc ggctgcgcct caacattcat gatcgtggcc accgtactgg
 780
 ctgttatccc actggcctgg ctgctcttcg cggccgtccg gcgcggcacc ggatcactat
 840
 tccacgcgtc gtggtggacc cactcgatgg atccctcctt cgacttggcc gagcagggcg
 900
 ccatccacgc tatcgtcgga acccttgaaa ttggccttat tacatcgatt atctcggtac
 960
 cgatcgctct gatgaccgcg atcttcttag tcgagtacgc ccgcggaact aagatcgcca
 1020
 aggtcattag cttcgcgcgc gacgtgctaa ccggtgtacc ttcaatcgtc gcggccctct
 1080
 tcgtcttcgc cgtagtcggt accaccttcg gtggcaccca atccgcgtgg gcctcctcgt
 1140
 tggeccatcat gatcctcatg gttccgacgg tgctgcgac aaccgaggaa atgctcaage
 1200
 tt
 1202

<210> 262

<211> 214

<212> PRT

<213> Homo sapiens

<400> 262

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Ala | Ser | Pro | Val | Ala | Phe | Val | Val | Asp | Leu | Leu | Ala | Ala | Val | Pro | Ser |
| 1 | | | | 5 | | | | 10 | | | | | 15 | | |
| Ile | Val | Phe | Gly | Leu | Trp | Gly | Gly | Ile | Val | Phe | Gly | Ser | Ser | Gly | Ile |
| | | | 20 | | | | | 25 | | | | | 30 | | |
| Ile | Asn | Gly | Tyr | Ala | Gly | Ala | Leu | Phe | Lys | Ala | Leu | Gly | Trp | Ile | Pro |
| | | | 35 | | | | 40 | | | | | 45 | | | |
| Ile | Phe | Ser | Glu | Asp | Pro | Ser | Trp | Ser | Ser | Ala | Thr | Gly | Thr | Val | Tyr |
| | | | 50 | | | | 55 | | | | 60 | | | | |
| Leu | Ala | Ser | Leu | Val | Leu | Ala | Ile | Met | Ile | Leu | Pro | Ile | Ile | Thr | Ala |
| 65 | | | | | 70 | | | | | 75 | | | | 80 | |
| Val | Ser | Arg | Asp | Val | Met | Pro | Arg | Thr | Pro | His | Asp | Gln | Val | Glu | Ala |
| | | | | 85 | | | | 90 | | | | | 95 | | |
| Ala | Leu | Ala | Leu | Gly | Ser | Thr | Arg | Trp | Glu | Val | Ile | Lys | Leu | Ala | Val |
| | | | 100 | | | | | 105 | | | | | 110 | | |
| Phe | Pro | His | Ser | Arg | Ser | Gly | Ile | Ile | Ser | Gly | Ser | Met | Leu | Gly | Leu |
| | | | 115 | | | | 120 | | | | | 125 | | | |
| Gly | Arg | Ala | Leu | Gly | Glu | Thr | Leu | Ala | Val | Thr | Leu | Ile | Leu | Gln | Thr |

```

      130              135              140
Met Ser Pro Met Ala Leu Lys Gln Asn Leu Asn Leu Ser Ile Phe Val
145              150              155              160
Gly Gly Glu Thr Phe Ala Ser Lys Ile Ala Gly Asn Phe Ser Glu Ala
      165              170              175
Ile Ser Asp Pro Thr Ser Leu Gly Ala Leu Val Ala Ser Ala Leu Ala
      180              185              190
Leu Phe Val Ile Thr Phe Val Val Asn Ala Thr Ala Arg Leu Ile Ala
      195              200              205
Ala Lys Gly Val Lys Arg
      210

```

<210> 263
 <211> 424
 <212> DNA
 <213> Homo sapiens

```

<400> 263
acgcgtgagt gctctgcgct ggaaacaacg gtgatagagc ccatccgcg tgaactttcc
60
gacgtggtgc tcgtgaacaa gctcgaaaag tatgtacgag aacgtacctc ggaagacgtt
120
gcgcacatgg aagaggatgc ggaccagacg ggcaacgaca tcctcacgac gatcctgctg
180
tcgaactggg atccactatt ggatatgacg acgcaggatc atgtgctggc catgcaaaag
240
gcttatatgg cctcgccatt ccgtgccaat ttggacctgg catacccatc ttcgacgcca
300
caggcccagt ccagccggc gatgccgcg tgggagacag ggacctcagc cagtagcatg
360
gcggatgctc gtgaatttgc gctgctgaag ctgtacctgc gtagcttgct gcagaagcac
420
gann
424

```

<210> 264
 <211> 99
 <212> PRT
 <213> Homo sapiens

```

<400> 264
Met Glu Glu Asp Ala Asp Gln Thr Gly Asn Asp Ile Leu Thr Thr Ile
1      5      10      15
Leu Leu Ser Asn Trp Asp Pro Leu Leu Asp Met Thr Thr Gln Asp His
      20      25      30
Val Leu Ala Met Gln Lys Ala Tyr Met Ala Ser Pro Phe Arg Ala Asn
      35      40      45
Leu Asp Leu Ala Tyr Pro Ser Ser Thr Pro Gln Ala Gln Ser Gln Pro
      50      55      60
Ala Met Pro Pro Trp Glu Thr Gly Thr Ser Ala Ser Ser Met Ala Asp
65      70      75      80
Ala Arg Glu Phe Ala Leu Leu Lys Leu Tyr Leu Arg Ser Leu Leu Gln
      85      90      95
Lys His Xaa

```


<210> 265
 <211> 360
 <212> DNA
 <213> Homo sapiens

<400> 265
 ncgtacggcc ctggcgctccg catggacgag ggataccatt ccggcatgac ggtgccgggt
 60
 gccttcgact ccctcatcgg caagctcacc atcactgggtg atagccgtga gcaagccctg
 120
 gctcgagctg cccgcgccct cgacgaaatc gtcacgacg gcatgccgac ggtcattccc
 180
 tttcaccagg cgggtggttca cgacccggct ttcactgccg ccgacggctg cttcggcgctc
 240
 tttaccgact ggatcgaaac cgagttcgac aacaagatcg agccatacac cgggtctctg
 300
 ggcgagtctg ccaattccga gcctcctcgt gaggtcgtcg tcgaggtaa cggtaaacgc
 360

<210> 266
 <211> 120
 <212> PRT
 <213> Homo sapiens

<400> 266
 Xaa Tyr Gly Pro Gly Val Arg Met Asp Glu Gly Tyr His Ser Gly Met
 1 5 10 15
 Thr Val Pro Gly Ala Phe Asp Ser Leu Ile Gly Lys Leu Ile Ile Thr
 20 25 30
 Gly Asp Ser Arg Glu Gln Ala Leu Ala Arg Ala Ala Arg Ala Leu Asp
 35 40 45
 Glu Ile Val Ile Asp Gly Met Pro Thr Val Ile Pro Phe His Gln Ala
 50 55 60
 Val Val His Asp Pro Ala Phe Thr Ala Ala Asp Gly Cys Phe Gly Val
 65 70 75 80
 Phe Thr Asp Trp Ile Glu Thr Glu Phe Asp Asn Lys Ile Glu Pro Tyr
 85 90 95
 Thr Gly Ser Leu Gly Glu Ser Ala Asn Ser Glu Pro Pro Arg Glu Val
 100 105 110
 Val Val Glu Val Asn Gly Lys Arg
 115 120

<210> 267
 <211> 471
 <212> DNA
 <213> Homo sapiens

<400> 267
 natcctcaac gtgtgttcag ttccacgcga aagatcatgt tcgtcatcgg atcgatgccg
 60
 ttaacgcac ctagtcaatc caccgatggc gaccctggca aaaaatacga ggtgacttgg
 120

ctagatctcg ggcaccttca ccctagtcgg ccgggactcg tcactatcac cacaactgtc
 180
 gatgatgacg tcatcacctc ttcccaggta aatgtcggca acctccaccg cggggatgaa
 240
 aaacttttcg aagctcgcga ttaccgccag attccgatgc ttgcatcacg tcatggctgg
 300
 acagctccat tcattggtga gaccggcgca gcccatgcca tcgaggatgc gatgggcatt
 360
 accatcccaa ctgcgtggc atggatacga accctgctcg ctgagttcag cagaatcacc
 420
 tcacacttca catttttgtc atgggtaggc catcactgtg atgatgccgg c
 471

<210> 268

<211> 157

<212> PRT

<213> Homo sapiens

<400> 268

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Xaa | Pro | Gln | Arg | Val | Phe | Ser | Ser | Thr | Arg | Lys | Ile | Met | Phe | Val | Ile |
| 1 | | | | 5 | | | | | 10 | | | | | 15 | |
| Gly | Ser | Met | Pro | Leu | Thr | His | Pro | Ser | Gln | Ser | Thr | Asp | Gly | Asp | Pro |
| | | | 20 | | | | | 25 | | | | | 30 | | |
| Gly | Lys | Lys | Tyr | Glu | Val | Thr | Trp | Leu | Asp | Leu | Gly | His | Leu | His | Pro |
| | | 35 | | | | 40 | | | | | | 45 | | | |
| Ser | Arg | Pro | Gly | Leu | Val | Thr | Ile | Thr | Thr | Thr | Val | Asp | Asp | Asp | Val |
| | 50 | | | | | 55 | | | | | 60 | | | | |
| Ile | Thr | Ser | Ser | Gln | Val | Asn | Val | Gly | Asn | Leu | His | Arg | Gly | Asp | Glu |
| 65 | | | | 70 | | | | | 75 | | | | | 80 | |
| Lys | Leu | Phe | Glu | Ala | Arg | Asp | Tyr | Arg | Gln | Ile | Pro | Met | Leu | Ala | Ser |
| | | | 85 | | | | | 90 | | | | | 95 | | |
| Arg | His | Gly | Trp | Thr | Ala | Pro | Phe | Ile | Gly | Glu | Thr | Gly | Ala | Ala | His |
| | | | 100 | | | | | 105 | | | | | 110 | | |
| Ala | Ile | Glu | Asp | Ala | Met | Gly | Ile | Thr | Ile | Pro | Thr | Arg | Val | Ala | Trp |
| | | 115 | | | | 120 | | | | | | 125 | | | |
| Ile | Arg | Thr | Leu | Leu | Ala | Glu | Phe | Ser | Arg | Ile | Thr | Ser | His | Phe | Thr |
| | 130 | | | | | 135 | | | | | 140 | | | | |
| Phe | Leu | Ser | Trp | Val | Gly | His | His | Cys | Asp | Asp | Ala | Gly | | | |
| 145 | | | | | 150 | | | | | 155 | | | | | |

<210> 269

<211> 387

<212> DNA

<213> Homo sapiens

<400> 269

acgcgtgtcg tggtttccaga aaaaaccaat aaattagagt ttatggtaga agtgattgct
 60
 gatatgacgg taatcaatcc atttgatttc tttgtggaaa gctacgcaga agactaccca
 120
 tttgcttatg acaaagctct taaaaaagag ttagaacctt atttacaggt ttctgaacct
 180
 tggttcgttac tcgacaaatg gctgtctggt gttgatcgtg aaaaaacacc gatcaatgat
 240

tttctagtcg caataaacag tcgccttgcc ggtgatattg gctatggat tcgcttagaa
 300
 ccgggcgttc agtcacctga agaaacgctc acattaatga aaggctcttg tcgcgatacc
 360
 tcgggggttat tgggtcaaact actacgc
 387

<210> 270

<211> 129

<212> PRT

<213> Homo sapiens

<400> 270

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Thr | Arg | Val | Val | Phe | Pro | Glu | Lys | Thr | Asn | Lys | Leu | Glu | Phe | Met | Val |
| 1 | | | | 5 | | | | | 10 | | | | | 15 | |
| Glu | Val | Ile | Ala | Asp | Met | Thr | Val | Ile | Asn | Pro | Phe | Asp | Phe | Phe | Val |
| | | | 20 | | | | | 25 | | | | | 30 | | |
| Glu | Ser | Tyr | Ala | Glu | Asp | Tyr | Pro | Phe | Ala | Tyr | Asp | Lys | Ala | Leu | Lys |
| | | 35 | | | | | 40 | | | | | 45 | | | |
| Lys | Glu | Leu | Glu | Pro | Tyr | Leu | Gln | Val | Ser | Glu | Pro | Cys | Ser | Leu | Leu |
| | | 50 | | | | 55 | | | | | 60 | | | | |
| Asp | Lys | Trp | Leu | Ser | Gly | Val | Asp | Arg | Glu | Lys | Thr | Pro | Ile | Asn | Asp |
| 65 | | | | | 70 | | | | | 75 | | | | 80 | |
| Phe | Leu | Val | Ala | Ile | Asn | Ser | Arg | Leu | Ala | Gly | Asp | Ile | Gly | Tyr | Gly |
| | | | 85 | | | | | | 90 | | | | | 95 | |
| Ile | Arg | Leu | Glu | Pro | Gly | Val | Gln | Ser | Pro | Glu | Glu | Thr | Leu | Thr | Leu |
| | | | 100 | | | | | 105 | | | | | 110 | | |
| Met | Lys | Gly | Ser | Cys | Arg | Asp | Thr | Ser | Gly | Leu | Leu | Val | Gln | Ile | Leu |
| | | 115 | | | | | 120 | | | | | | 125 | | |

Arg

<210> 271

<211> 443

<212> DNA

<213> Homo sapiens

<400> 271

gccggcacca acggaaagtc ctctaccgcg cgcattggctg attcgctttt gcgtgccttc
 60
 caccgccgag tgggttttgg aaccagccca cacctgcagc gcgttactga gcgcatcggc
 120
 attgatggcc agcccattca cccgcgcgat tatgtacgca tctggcacga gattaagcca
 180
 tttgtggaaa tggtcgatgc cgaatcggac gtgcctatgt ctaagtccga ggtcttcgtg
 240
 ggctgtcct atgctgcgtt tgccgacgcc cccggggacg tcgctgtcgt cgaagtcggc
 300
 cttggcggac gttgggacgc taccaatgtg gtcaacgcgg atgtctctgt cattaccccg
 360
 gtgggcatgg accacacgga ttacctgggg gagacgatca ctgaaatcgc aggcgagaaa
 420
 gctggcatta ttaagccacg cgt
 443

<210> 272
 <211> 147
 <212> PRT
 <213> Homo sapiens

<400> 272
 Ala Gly Thr Asn Gly Lys Ser Ser Thr Ala Arg Met Val Asp Ser Leu
 1 5 10 15
 Leu Arg Ala Phe His Arg Arg Val Gly Leu Val Thr Ser Pro His Leu
 20 25 30
 Gln Arg Val Thr Glu Arg Ile Gly Ile Asp Gly Gln Pro Ile His Pro
 35 40 45
 Arg Asp Tyr Val Arg Ile Trp His Glu Ile Lys Pro Phe Val Glu Met
 50 55 60
 Val Asp Ala Glu Ser Asp Val Pro Met Ser Lys Phe Glu Val Phe Val
 65 70 75 80
 Gly Leu Ser Tyr Ala Ala Phe Ala Asp Ala Pro Gly Asp Val Ala Val
 85 90 95
 Val Glu Val Gly Leu Gly Gly Arg Trp Asp Ala Thr Asn Val Val Asn
 100 105 110
 Ala Asp Val Ser Val Ile Thr Pro Val Gly Met Asp His Thr Asp Tyr
 115 120 125
 Leu Gly Glu Thr Ile Thr Glu Ile Ala Gly Glu Lys Ala Gly Ile Ile
 130 135 140
 Lys Pro Arg
 145

<210> 273
 <211> 864
 <212> DNA
 <213> Homo sapiens

<400> 273
 caaagtaaga ctgcttcaaa ttttgtgttc tgctctgcag ctgctcccc cctgctgtcg
 60
 aagagaagcc aaagcccccc cccccacct caaaggctcg gaagtctggc atccctactt
 120
 ccgagcctgg atcccagtaa ggatcttgcc ctccctgcaa caccgagtgc cttagacagc
 180
 tgctgctga gaactggcct ccagccggtg tcctcattcc atggggctcc ctgctgactg
 240
 catttctga tctgggatga tgtttaccag cccaaaacca gtcattgttct tccaaaagct
 300
 tctctttgat agaattttga ggccatgcca cctcccttcc agtccacatg gaattccaga
 360
 atcagtcaca gcctctgatt ttttccaaga agagattgcc ttcaccattg ttaaatgtca
 420
 gcctgtacgg cagagacatg gtggtctgca caagcctgga caagttcttc catattgatg
 480
 gtgggagcaa cccctgtaat ctactccttg gaaggatttt ttgctttgct tatgaaaagc
 540
 tgtgcttgag acttaggtac ttttctcacg tggacacact gatcccatcc catattgcat
 600

ctttgaagag atggatatca agtacacttt ggtagctgaa ataatcatat ctttctgatg
 660
 tctattgtat ctcttttgag gaaaagaaca cacattttta atggagattg gctgctttca
 720
 ggtatgtgtg tctatcattg aaagagcatg gactcaaaca tcagccctga gttcttgagt
 780
 ccaccaact cccatcttct tgtggcacag gaaagctgcc ctctccctct cccaccacac
 840
 tcctgactaa tgcccttcac gcgt
 864

<210> 274
 <211> 116
 <212> PRT
 <213> Homo sapiens

<400> 274
 Met Trp Thr Gly Arg Glu Val Ala Trp Pro Gln Asn Ser Ile Lys Glu
 1 5 10 15
 Lys Leu Leu Glu His Asp Trp Phe Trp Ala Gly Lys His His Pro
 20 25 30
 Arg Ser Gly Asn Ala Val Ser Arg Glu Pro His Gly Met Arg Thr Pro
 35 40 45
 Ala Gly Gly Gln Phe Ser Gly Ser Ser Cys Leu Arg His Ser Val Leu
 50 55 60
 Gln Gly Gly Gln Asp Pro Tyr Trp Asp Pro Gly Ser Glu Val Gly Met
 65 70 75 80
 Pro Asp Phe Arg Ala Phe Glu Val Gly Gly Gly Phe Gly Phe Ser
 85 90 95
 Ser Thr Ala Gly Ser Glu Leu Gln Ser Arg Thr Gln Asn Leu Lys
 100 105 110
 Gln Ser Tyr Phe
 115

<210> 275
 <211> 911
 <212> DNA
 <213> Homo sapiens

<400> 275
 naaatttaaa ggaacctccc ttctataacg gagagtattt attgcagctt tcctttctgt
 60
 ttattttcag gaatgaaagg aattaccag ccttctgctt ttatacctac agctgaaagt
 120
 aattcctttc agcctcaggt gaagactttg ccattctcaa ttgatgctaa acagcagttg
 180
 caacggaaaa tccagaagaa gcagcaagaa cagaaactac aatccccctt gccaggagaa
 240
 tctgcagcaa aaaagtcaga aagtgtaca agcaatggag tgactaatct tcctaattgga
 300
 aatccttcaa tcctttctcc tcaacctatt ggtatcggtg tggcagctgt ccctagtccc
 360
 attccggtcc agcggactag gcaattggta acttcaccga gtccaatgag ttcttctnga
 420

cggcaaagtt cttccctca atgtacaggt ggtcactcag cacatgcagt ctgtgaaaca
 480
 ggcaccaaag actccccaga acgttccagc agtcctgggtg ggaatcggtc tgcccggcac
 540
 cgttaccctc agatcttacc caaaccagcg aacaccagtg cactcaccat tcgctctcca
 600
 actactgtcc tctttactag tagtcccatc aaaactgctg ttgtaccgcg ttcacacatg
 660
 agttctctaa atgtggtgaa aatgacaaca atatccctca caccagcaa cagtaacacc
 720
 cctcttaaac attctgcctc agtcagcagt gctacaggaa caacagaaga atcaaggagt
 780
 gttccacaga tcaagaatgg ttctgtcgtg tcgcttcagt ctctggggtc caggagcagc
 840
 agtgcggggg gaacatctgc tgtggaagtc aaagtggaac ccgaaacatc atcagatgag
 900
 catcctgtac a
 911

<210> 276

<211> 279

<212> PRT

<213> Homo sapiens

<400> 276

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Lys | Gly | Ile | Thr | Gln | Pro | Ser | Ala | Phe | Ile | Pro | Thr | Ala | Glu | Ser |
| 1 | | | 5 | | | | | | 10 | | | | | 15 | |
| Asn | Ser | Phe | Gln | Pro | Gln | Val | Lys | Thr | Leu | Pro | Ser | Pro | Ile | Asp | Ala |
| | | | 20 | | | | | | 25 | | | | 30 | | |
| Lys | Gln | Gln | Leu | Gln | Arg | Lys | Ile | Gln | Lys | Lys | Gln | Gln | Glu | Gln | Lys |
| | | | 35 | | | | 40 | | | | | | 45 | | |
| Leu | Gln | Ser | Pro | Leu | Pro | Gly | Glu | Ser | Ala | Ala | Lys | Lys | Ser | Glu | Ser |
| | | | 50 | | | | 55 | | | | | 60 | | | |
| Ala | Thr | Ser | Asn | Gly | Val | Thr | Asn | Leu | Pro | Asn | Gly | Asn | Pro | Ser | Ile |
| 65 | | | | | 70 | | | | | 75 | | | | 80 | |
| Leu | Ser | Pro | Gln | Pro | Ile | Gly | Ile | Val | Val | Ala | Ala | Val | Pro | Ser | Pro |
| | | | 85 | | | | | 90 | | | | | | 95 | |
| Ile | Pro | Val | Gln | Arg | Thr | Arg | Gln | Leu | Val | Thr | Ser | Pro | Ser | Pro | Met |
| | | | 100 | | | | | 105 | | | | | 110 | | |
| Ser | Ser | Ser | Xaa | Arg | Gln | Ser | Ser | Ser | Pro | Gln | Cys | Thr | Gly | Gly | His |
| | | | 115 | | | | 120 | | | | | 125 | | | |
| Ser | Ala | His | Ala | Val | Cys | Glu | Thr | Gly | Thr | Lys | Asp | Ser | Pro | Glu | Arg |
| | | | 130 | | | | 135 | | | | 140 | | | | |
| Ser | Ser | Ser | Pro | Gly | Gly | Asn | Arg | Ser | Ala | Arg | His | Arg | Tyr | Pro | Gln |
| 145 | | | | | 150 | | | | | 155 | | | | 160 | |
| Ile | Leu | Pro | Lys | Pro | Ala | Asn | Thr | Ser | Ala | Leu | Thr | Ile | Arg | Ser | Pro |
| | | | 165 | | | | | | 170 | | | | | 175 | |
| Thr | Thr | Val | Leu | Phe | Thr | Ser | Ser | Pro | Ile | Lys | Thr | Ala | Val | Val | Pro |
| | | | 180 | | | | | 185 | | | | | 190 | | |
| Ala | Ser | His | Met | Ser | Ser | Leu | Asn | Val | Val | Lys | Met | Thr | Thr | Ile | Ser |
| | | | 195 | | | | 200 | | | | | 205 | | | |
| Leu | Thr | Pro | Ser | Asn | Ser | Asn | Thr | Pro | Leu | Lys | His | Ser | Ala | Ser | Val |
| | | | 210 | | | 215 | | | | | 220 | | | | |
| Ser | Ser | Ala | Thr | Gly | Thr | Thr | Glu | Glu | Ser | Arg | Ser | Val | Pro | Gln | Ile |

| | | | | | | | | | | | | | | | |
|-----------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| <400> 278 | | | | | | | | | | | | | | | |
| Met | Ser | Glu | Val | Pro | Asp | Glu | Leu | Val | Val | Leu | Arg | Gly | Ala | Ile | Asp |
| 1 | | | | 5 | | | | | 10 | | | | | 15 | |
| Asn | Met | Asp | Ala | Leu | Ile | His | Leu | Leu | Ala | Glu | Arg | Phe | Arg | Ile | |
| | | | 20 | | | | 25 | | | | | | 30 | | |
| Thr | Arg | Glu | Val | Gly | Arg | Leu | Lys | Ala | Glu | Cys | Gly | Leu | Pro | Pro | Ala |
| | | 35 | | | | 40 | | | | | | 45 | | | |
| Asp | Pro | Ala | Arg | Glu | Ala | Glu | Gln | Ile | Ala | Arg | Leu | Arg | Gln | Leu | Ala |
| | | 50 | | | | 55 | | | | | 60 | | | | |
| Val | Glu | Ser | Asn | Leu | Asp | Pro | Glu | Phe | Ala | Gln | Lys | Val | Ile | Thr | Phe |
| 65 | | | | | 70 | | | | | 75 | | | | | 80 |
| Ile | Val | Ala | Glu | Val | Val | Arg | His | His | Glu | Ala | Ile | Ala | Asp | Asp | Ser |

85 90 95
 Gly Asp Asp Ser Gly Val Ala Asp Thr Gly Glu Ala Asp Val Pro Gly
 100 105 110
 Ser Gly Ser
 115

<210> 279
 <211> 348
 <212> DNA
 <213> Homo sapiens

<400> 279
 cgggaggtca cacaagcatt caaaccatag cagatggtaa atgttatggt atgtgtatatt
 60
 taccacaatc cttaaaaaga aaagaaagaa aggcatatgg aacccttagt tacctctcat
 120
 ccagcttcaa aattgtcagt gcatgggtcaa tcttgtctta tctgccctc acccaccctt
 180
 ttccagaaag aagaccaga ggattccaca tctgcctgga aaccacgacc agtctcgact
 240
 ggaagtgtgt gttaatgttg catgtattca taaaacctct aggcatttct agtgcctcctc
 300
 agaatttttc caaattcagg caaacacaga aattacttcc aaaaattt
 348

<210> 280
 <211> 99
 <212> PRT
 <213> Homo sapiens

<400> 280
 Met Cys Ile Leu Pro Gln Ser Leu Lys Arg Lys Glu Arg Lys Ala Tyr
 1 5 10 15
 Gly Thr Pro Ser Tyr Leu Ser Ser Ser Phe Lys Ile Val Ser Ala Trp
 20 25 30
 Ser Ile Leu Ser Tyr Leu Pro Leu Thr His Pro Phe Pro Glu Arg Arg
 35 40 45
 Pro Arg Gly Phe His Ile Cys Leu Glu Thr Thr Thr Ser Leu Asp Trp
 50 55 60
 Lys Leu Leu Leu Met Leu His Val Phe Ile Lys Pro Leu Gly Ile Ser
 65 70 75 80
 Ser Val Pro Gln Asn Phe Ser Lys Phe Arg Gln Thr Gln Lys Leu Leu
 85 90 95
 Pro Lys Ile

<210> 281
 <211> 384
 <212> DNA
 <213> Homo sapiens

<400> 281
 agatctgcgc agatcgataa tggattaaag actcttgacg ctggagtcac cgagatgaac
 60

aacaagggtgt tgggggcaac gaaggctgtc ggtgattcca ccactaccgt caaccagggtg
 120
 aattctgcgt taggaantgc cgactcagcg gcagagaaga cgtcgagcgc cgttactcag
 180
 acgcgcgtgg gtgcccaggc gattaccggc gctgctcaaa atgtcatggc tgattcccaa
 240
 gctgtcaact cagccatggt tccgcttatt aataacgtga caaagaatct tcctaccttg
 300
 caaaaacagg ccaggaatct cgtgtcagtg aacggtaccc tgcagaaccc caacggtgat
 360
 tctgtcatta agattcaaca gacc
 384

<210> 282
 <211> 110
 <212> PRT
 <213> Homo sapiens

<400> 282
 Met Asn Asn Lys Val Leu Gly Ala Thr Lys Ala Val Gly Asp Ser Thr
 1 5 10 15
 Thr Thr Val Asn Gln Val Asn Ser Ala Leu Gly Xaa Ala Asp Ser Ala
 20 25 30
 Ala Glu Lys Thr Ser Ser Ala Val Thr Gln Thr Arg Val Gly Ala Gln
 35 40 45
 Ala Ile Thr Gly Ala Ala Gln Asn Val Met Ala Asp Ser Gln Ala Val
 50 55 60
 Asn Ser Ala Met Val Pro Leu Ile Asn Asn Val Thr Lys Asn Leu Pro
 65 70 75 80
 Thr Leu Gln Lys Gln Ala Arg Asn Leu Val Ser Val Asn Gly Thr Leu
 85 90 95
 Gln Asn Pro Asn Gly Asp Ser Val Ile Lys Ile Gln Gln Thr
 100 105 110

<210> 283
 <211> 426
 <212> DNA
 <213> Homo sapiens

<400> 283
 cgcgtagacc aatgtgagac ggccgtcacc aagggcatgc gcgacaagtc gggttggtagc
 60
 ggaccggata ttgtgcgtcg cgagctgcgc catgtcgtga cgagcggcac gattgtcgat
 120
 ggaagcgtac tggctgacga attgagcagc tactgcatga gtatcaagga gcacgtccgc
 180
 tctgatggcc tatccgagtt tggcatctgc accctcgacg ccgccaccgc cgagttccga
 240
 tacatgacat tcgtcgacga tgccgtgctg tcacaactcg agacattgct gcgttctcta
 300
 cgcacaaagg aagtcttgca tgaaaaaggg gtcattgttc cttccacgct gcgcttgatc
 360
 cgcaacgcgg tgcccaccac ctgcctaaatt accatgctca agcctgatac cgaattgtcg
 420

gagaga
426

<210> 284
<211> 142
<212> PRT
<213> Homo sapiens

<400> 284
Arg Val Asp Gln Cys Glu Thr Ala Val Thr Lys Gly Met Arg Asp Lys
1 5 10 15
Ser Val Gly Ser Gly Pro Asp Ile Val Arg Arg Glu Leu Arg His Val
20 25 30
Val Thr Ser Gly Thr Ile Val Asp Gly Ser Val Leu Ala Asp Glu Leu
35 40 45
Ser Ser Tyr Cys Met Ser Ile Lys Glu His Val Arg Ser Asp Gly Leu
50 55 60
Ser Glu Phe Gly Ile Cys Thr Leu Asp Ala Ala Thr Ala Glu Phe Arg
65 70 75 80
Tyr Met Thr Phe Val Asp Asp Ala Val Leu Ser Gln Leu Glu Thr Leu
85 90 95
Leu Arg Ser Leu Arg Ile Lys Glu Val Leu His Glu Lys Gly Val Met
100 105 110
Leu Pro Ser Thr Leu Arg Leu Ile Arg Asn Ala Val Pro Thr Thr Cys
115 120 125
Gln Ile Thr Met Leu Lys Pro Asp Thr Glu Leu Ser Glu Arg
130 135 140

<210> 285
<211> 345
<212> DNA
<213> Homo sapiens

<400> 285
acgcgtgcag tcccttaccg acatgctggc agatgagctc gacggcagcc gcttcaccgg
60
cgatttctca gaaatctaca aacgtcagaa ctcgatcttc ggcatgtaa ggaataactt
120
ttacaaaaaa ggataccgca tcatcaacgt agcgaatggt gtattgcgca agatttcact
180
ggtaagcgca ggcaatgcag acaatgtgaa aggtcaggcc ctgttcttcc gcggtgtggc
240
gcatttcgaa ctcggtgcgtt tgtttgcaca accctgggggt tatacttcgg acaattcaca
300
ctacggcatc ccgctccgca atgaaatcgt aattggttct attcn
345

<210> 286
<211> 107
<212> PRT
<213> Homo sapiens

<400> 286
Met Leu Ala Asp Glu Leu Asp Gly Ser Arg Phe Thr Gly Asp Phe Ser

```

      1           5           10           15
Glu Ile Tyr Lys Arg Gln Asn Ser Ile Phe Gly Asp Val Arg Asn Asn
      20           25           30
Phe Tyr Lys Lys Gly Tyr Arg Ile Ile Asn Val Ala Asn Gly Val Leu
      35           40           45
Arg Lys Ile Ser Leu Val Ser Ala Gly Asn Ala Asp Asn Val Lys Gly
      50           55           60
Gln Ala Leu Phe Phe Arg Gly Val Ala His Phe Glu Leu Val Arg Leu
      65           70           75           80
Phe Ala Gln Pro Trp Gly Tyr Thr Ser Asp Asn Ser His Tyr Gly Ile
      85           90           95
Pro Leu Arg Asn Glu Ile Val Ile Gly Ser Ile
      100           105

```

<210> 287

<211> 1379

<212> DNA

<213> Homo sapiens

<400> 287

```

nnttaactgc ccctttgcag tctttattct gggacattag cactgtctgg ttatcttgct
60
tcagttgagg gattcgggac aatagcagtg ctgatggtaa tggtggcgat ttccctgttt
120
gttttgcagg tcacggccag gggctttggg ccgctgttac agtttgcta cactgccaaag
180
ctgttactca gcagagaaaa catccgcgag gtcacccgct gtgctgagtt cctgcgcag
240
cacaacctgg aggactcctg cttcagcttc ctgcagaccc agctcctgaa cagtgaggat
300
ggcctgtttg tgtgccggaa ggatgctgag tgccagcgcc cacacgagga ctgcgagaac
360
tctgcaggag aggaggagga tgaagaggag gagacgatgg attcagagac ggccaagatg
420
gcttgcccca gggaccagat gcttcagag cccatcagct ttgaggccgc cgccatcccc
480
gtagcagaga aggaagaagc cctgctgccc gagcctgacg tgcccacaga caccaaggag
540
agctcagaaa aggacgcgtt aacgcagtac ccagataca agaaatacca gcttgcatgt
600
accaagaatg tctataatgc atcatcacac agtacctcag gttttgcaag cacattccgg
660
gaagataact ctagcaacag cctcaagccg gggcttgcca gggggcagat taaaagttag
720
ccgcccagtg aagagaatga ggaagagagc atcacgctct gcctgtctgg agatgagcct
780
gacgccaagg acagagcggg ggatgtcgag atggaccgga aacagcccag ccctgcccct
840
acccccacgg ccccagctgg ggccgctgct ctggagagat ccaggagcgt ggctcgcgcc
900
tcctgcttaa ggtctctgtt cagcataacg aaaagtgtgg agctgtctgg cctgcccagt
960
acatctcagc agcactttgc caggagtcca gcctgccctt ttgacaaggg gatcactcag
1020

```

ggtgacctta aaactgacta cacccttttc acaggaatt atggacagcc ccacgtgggc
 1080
 cagaaggagg tgtccaactt caccatgggg tcgcccctca gggggcctgg gttggaggct
 1140
 ctctgtaaac aggagggaga gctggaccgg aggagcgtga tcttctctc cagcgcttgt
 1200
 gaccaagtga gcacctcggt gcattcttat tctgggggtga gcagtttga caaagacctc
 1260
 tctgagccgg tgccaaaggg tctgtgggtg ggagccggcc agtccctccc cagctcgcag
 1320
 gcctactccc acgggtgggt gatggccgac cacttgccag gaaggatgcg gcccaacac
 1379

<210> 288

<211> 428

<212> PRT

<213> Homo sapiens

<400> 288

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Val | Met | Leu | Ala | Ile | Ser | Leu | Phe | Val | Leu | Gln | Val | Thr | Ala | Arg |
| 1 | | | | 5 | | | | | 10 | | | | | 15 | |
| Gly | Phe | Gly | Pro | Leu | Leu | Gln | Phe | Ala | Tyr | Thr | Ala | Lys | Leu | Leu | Leu |
| | | | 20 | | | | | 25 | | | | 30 | | | |
| Ser | Arg | Glu | Asn | Ile | Arg | Glu | Val | Ile | Arg | Cys | Ala | Glu | Phe | Leu | Arg |
| | | 35 | | | | 40 | | | | | 45 | | | | |
| Met | His | Asn | Leu | Glu | Asp | Ser | Cys | Phe | Ser | Phe | Leu | Gln | Thr | Gln | Leu |
| | 50 | | | | 55 | | | | | 60 | | | | | |
| Leu | Asn | Ser | Glu | Asp | Gly | Leu | Phe | Val | Cys | Arg | Lys | Asp | Ala | Ala | Cys |
| 65 | | | | 70 | | | | 75 | | | | | | 80 | |
| Gln | Arg | Pro | His | Glu | Asp | Cys | Glu | Asn | Ser | Ala | Gly | Glu | Glu | Glu | Asp |
| | | | 85 | | | | 90 | | | | | 95 | | | |
| Glu | Glu | Glu | Glu | Thr | Met | Asp | Ser | Glu | Thr | Ala | Lys | Met | Ala | Cys | Pro |
| | | | 100 | | | | 105 | | | | | 110 | | | |
| Arg | Asp | Gln | Met | Leu | Pro | Glu | Pro | Ile | Ser | Phe | Glu | Ala | Ala | Ala | Ile |
| | 115 | | | | | 120 | | | | | 125 | | | | |
| Pro | Val | Ala | Glu | Lys | Glu | Glu | Ala | Leu | Leu | Pro | Glu | Pro | Asp | Val | Pro |
| | 130 | | | | 135 | | | | | 140 | | | | | |
| Thr | Asp | Thr | Lys | Glu | Ser | Ser | Glu | Lys | Asp | Ala | Leu | Thr | Gln | Tyr | Pro |
| 145 | | | | 150 | | | | | 155 | | | | | 160 | |
| Arg | Tyr | Lys | Lys | Tyr | Gln | Leu | Ala | Cys | Thr | Lys | Asn | Val | Tyr | Asn | Ala |
| | | | 165 | | | | 170 | | | | | | 175 | | |
| Ser | Ser | His | Ser | Thr | Ser | Gly | Phe | Ala | Ser | Thr | Phe | Arg | Glu | Asp | Asn |
| | | 180 | | | | 185 | | | | | | 190 | | | |
| Ser | Ser | Asn | Ser | Leu | Lys | Pro | Gly | Leu | Ala | Arg | Gly | Gln | Ile | Lys | Ser |
| | | 195 | | | | 200 | | | | | 205 | | | | |
| Glu | Pro | Pro | Ser | Glu | Glu | Asn | Glu | Glu | Glu | Ser | Ile | Thr | Leu | Cys | Leu |
| | 210 | | | | 215 | | | | | 220 | | | | | |
| Ser | Gly | Asp | Glu | Pro | Asp | Ala | Lys | Asp | Arg | Ala | Gly | Asp | Val | Glu | Met |
| 225 | | | | 230 | | | | | 235 | | | | | 240 | |
| Asp | Arg | Lys | Gln | Pro | Ser | Pro | Ala | Pro | Thr | Pro | Thr | Ala | Pro | Ala | Gly |
| | | | 245 | | | | 250 | | | | | | 255 | | |
| Ala | Ala | Cys | Leu | Glu | Arg | Ser | Arg | Ser | Val | Ala | Ser | Pro | Ser | Cys | Leu |
| | | 260 | | | | 265 | | | | | | 270 | | | |
| Arg | Ser | Leu | Phe | Ser | Ile | Thr | Lys | Ser | Val | Glu | Leu | Ser | Gly | Leu | Pro |

| | | |
|---|-----|-----|
| 275 | 280 | 285 |
| Ser Thr Ser Gln Gln His Phe Ala Arg Ser Pro Ala Cys Pro Phe Asp | | |
| 290 | 295 | 300 |
| Lys Gly Ile Thr Gln Gly Asp Leu Lys Thr Asp Tyr Thr Pro Phe Thr | | |
| 305 | 310 | 315 |
| Gly Asn Tyr Gly Gln Pro His Val Gly Gln Lys Glu Val Ser Asn Phe | | |
| | 325 | 330 |
| Thr Met Gly Ser Pro Leu Arg Gly Pro Gly Leu Glu Ala Leu Cys Lys | | |
| | 340 | 345 |
| Gln Glu Gly Glu Leu Asp Arg Arg Ser Val Ile Phe Ser Ser Ser Ala | | |
| | 355 | 360 |
| Cys Asp Gln Val Ser Thr Ser Val His Ser Tyr Ser Gly Val Ser Ser | | |
| | 370 | 375 |
| Leu Asp Lys Asp Leu Ser Glu Pro Val Pro Lys Gly Leu Trp Val Gly | | |
| 385 | 390 | 395 |
| Ala Gly Gln Ser Leu Pro Ser Ser Gln Ala Tyr Ser His Gly Gly Leu | | |
| | 405 | 410 |
| Met Ala Asp His Leu Pro Gly Arg Met Arg Pro Asn | | |
| | 420 | 425 |

<210> 289
 <211> 822
 <212> DNA
 <213> Homo sapiens

<400> 289
 ngcattaccg ggctgaagac ggggtgctcat gacctcaacg atataggcta ttgctagaac
 60
 cagcccgcc cagccgcgc aaagcgcaga cagggcacca ggaggggtca catggctgat
 120
 agcaagtga aggcgaagga cgagcgcact gccgatgaga tcaggcggga tattgcagcg
 180
 acccgtgctt gcctggcagc cggggtggag aacctcgtgg aggaggtgca tccggcaacc
 240
 ctcaagcgtg aagcatctga tcgtgcccgt gattttgtgc aggggtgagtt tgatcaggtc
 300
 aagagccagg tcaaagatga gaaatggtgg cgcgtgcagc ggatcgcgat ggccgcagga
 360
 gtgctcgctg ccggcgtcgt cagcattatt gtgctgcgcg cgatagtcgg tcgcgcaacg
 420
 ggcgctaccg ctgcgcgcaa gcttgagaag ctgcagcttt ctcaggcgaa gcgggttcga
 480
 aaagatgcc aagcagcgtag taaggaagat gaaaaggcag ccaagaaaaa tgccaagctc
 540
 ggcaagaaga acgctaagaa gtacggcaag ctcgataccg atgactcgtc ggtaagcaac
 600
 cttgccgaga aaatgctcaa acaggccgcc gtgctgcgtg cacaggcggc tgccggggcg
 660
 tgagaacagt gccgcctagc aaacagcggc cacagcgcaa aacaggtttg gctccgaccc
 720
 atggtggacc ggagccaaac tgtgttaccg catcatttga taccgccagc agccaggcct
 780
 gcgacaatgc gacgctggaa taccagcacc atgatgacta gt
 822

<210> 290
 <211> 183
 <212> PRT
 <213> Homo sapiens

<400> 290
 Met Ala Asp Ser Lys Ser Lys Ala Lys Asp Glu Arg Thr Ala Asp Glu
 1 5 10 15
 Ile Arg Arg Asp Ile Ala Ala Thr Arg Ala Cys Leu Ala Ala Gly Val
 20 25 30
 Glu Asn Leu Val Glu Glu Val His Pro Ala Thr Leu Lys Arg Glu Ala
 35 40 45
 Ser Asp Arg Ala Arg Asp Phe Val Gln Gly Glu Phe Asp Gln Val Lys
 50 55 60
 Ser Gln Val Lys Asp Glu Lys Trp Trp Arg Val Gln Arg Ile Ala Met
 65 70 75 80
 Ala Ala Gly Val Leu Ala Ala Gly Val Val Ser Ile Ile Val Leu Arg
 85 90 95
 Ala Ile Val Gly Arg Ala Thr Gly Ala Thr Ala Arg Arg Lys Leu Glu
 100 105 110
 Lys Leu Gln Leu Ser Gln Ala Lys Arg Val Arg Lys Asp Ala Lys Gln
 115 120 125
 Arg Ser Lys Glu Asp Glu Lys Ala Ala Lys Lys Asn Ala Lys Leu Gly
 130 135 140
 Lys Lys Asn Ala Lys Lys Tyr Gly Lys Leu Asp Thr Asp Asp Ser Ser
 145 150 155 160
 Val Ser Asn Leu Ala Glu Lys Met Leu Lys Gln Ala Ala Val Leu Arg
 165 170 175
 Ala Gln Ala Ala Ala Gly Ala
 180

<210> 291
 <211> 351
 <212> DNA
 <213> Homo sapiens

<400> 291
 ctccacgccg acaagactta cgacgggcgt cgctgccggg ctgagtgccg ggcccgcctcc
 60
 atcacccccc gcatcgctcg ccgcggcggtg gagaccagcg agcgcttggg ccggtatcgc
 120
 tgggtcgctg agcgcacctt cgctgggctc aaccgctttc ggcgccctcgc catccgctac
 180
 gagcggcggtg ctgacatcca cgaagccttc gtgacccctg gctgcgcctt catctgcctc
 240
 aaccagatca gacggttttg ttaggtgctg taaagggaga atggctgcag ctgggctatc
 300
 tgctccctcg tcaaccagaa acaggctgct catcctcact caacaacgcg t
 351

<210> 292
 <211> 87
 <212> PRT

<213> Homo sapiens

<400> 292

```

Leu His Ala Asp Lys Thr Tyr Asp Gly Arg Arg Cys Arg Ala Glu Cys
 1           5           10           15
Arg Ala Arg Ser Ile Thr Pro Arg Ile Ala Arg Arg Gly Val Glu Thr
          20           25           30
Ser Glu Arg Leu Gly Arg Tyr Arg Trp Val Val Glu Arg Thr Phe Ala
          35           40           45
Trp Leu Asn Arg Phe Arg Arg Leu Ala Ile Arg Tyr Glu Arg Arg Ala
          50           55           60
Asp Ile His Glu Ala Phe Val Ile Leu Gly Cys Ala Leu Ile Cys Leu
65           70           75           80
Asn Gln Ile Arg Arg Phe Cys
          85

```

<210> 293

<211> 716

<212> DNA

<213> Homo sapiens

<400> 293

```

nncttcacca caccggccat caacgcacct cctcgtgata acttgacett ctgccgaacc
60
ggttaatcag tttagtggcg aggcattgaca cggtgacgag tcagctgtgg tacatgtgcg
120
gaacactcac aatgccacgg cggcatgttg ctgtcgggtca cgacccttat ggtgatcgct
180
gtgagaaccc gaacggcaga tgcgattctg gcggcactgg atctgaacag gtttaagggt
240
gcgaagactt tcgatgttcc agtgtgcgtc atagctgggtg ccgggacagg taaaactcgt
300
gctgtcactc atcgcattgc ctacgggtgca gcgacaggca agcttgatcc gcgtcgtacc
360
ctcgcgggtca cttttacgac taaggcagct ggcacgatga gaggtcgact cgccgatctg
420
gggggtgttg gtgtgcaggc tcgcactatt cattctgcgg cgttgcggca gatcaagttt
480
ttctggcctc gtgcatataa ctgtgagttg ccaccggtga gtgattctcg tttctcgatg
540
gtggcggaga cgaccatcg cattgggtctg ggcaatgaca aggcgctgct gcgcgacttg
600
tccgcccaga tctcgtgggc gaagggtctca aatgtgccga ctgatcaata cgcattccctg
660
gctagggcgg aaggtcgggt ggtggcggga gtttcggcaa ctgacgtagg acgcgt
716

```

<210> 294

<211> 190

<212> PRT

<213> Homo sapiens

<400> 294

```

Met Leu Leu Ser Val Thr Thr Leu Met Val Ile Ala Val Arg Thr Arg

```

```

      1           5           10           15
Thr Ala Asp Ala Ile Leu Ala Ala Leu Asp Leu Asn Arg Phe Lys Val
      20           25           30
Ala Lys Thr Phe Asp Val Pro Val Cys Val Ile Ala Gly Ala Gly Thr
      35           40           45
Gly Lys Thr Arg Ala Val Thr His Arg Ile Ala Tyr Gly Ala Ala Thr
      50           55           60
Gly Lys Leu Asp Pro Arg Arg Thr Leu Ala Val Thr Phe Thr Thr Lys
      65           70           75           80
Ala Ala Gly Thr Met Arg Gly Arg Leu Ala Asp Leu Gly Val Val Gly
      85           90           95
Val Gln Ala Arg Thr Ile His Ser Ala Ala Leu Arg Gln Ile Lys Phe
      100          105          110
Phe Trp Pro Arg Ala Tyr Asn Cys Glu Leu Pro Pro Val Ser Asp Ser
      115          120          125
Arg Phe Ser Met Val Ala Glu Thr Thr His Arg Ile Gly Leu Gly Asn
      130          135          140
Asp Lys Ala Leu Leu Arg Asp Leu Ser Ala Glu Ile Ser Trp Ala Lys
      145          150          155          160
Val Ser Asn Val Pro Thr Asp Gln Tyr Ala Ser Leu Ala Arg Ala Glu
      165          170          175
Gly Arg Val Val Ala Gly Val Ser Ala Thr Asp Val Gly Arg
      180          185          190

```

<210> 295

<211> 417

<212> DNA

<213> Homo sapiens

<400> 295

```

ttcatatcag gcagtacccg agtccatgcg atcaacaacg tcagcgtatc tttcacccat
60
tctggagtgc accttctcat gggagaaaagc ggatcaggaa aaagcaccct catcaatctc
120
ctagctggtc tggatacccc agattcgggg tccgtctacg cagaaggcgt caccgtatct
180
gatcagagcg aggcgagcag agcccaatct cgattacgcc acatcgccgt catcttccag
240
gacgacaacc tcatcgctga gttgaccaat accgagaata ttgcgctacc cctgtgggcg
300
cagggcacat cgaagtccga tgccactgaa atcgcccacg aagccatgcg aaaactagga
360
atcgagtcac tgggcagacg ctaccccggc gaggtctcgg gtggccaacg gcaacgc
417

```

<210> 296

<211> 139

<212> PRT

<213> Homo sapiens

<400> 296

```

Phe Ile Ser Gly Ser Thr Arg Val His Ala Ile Asn Asn Val Ser Val
1           5           10           15
Ser Phe Thr His Ser Gly Val His Leu Leu Met Gly Glu Ser Gly Ser

```



```

      20      25      30
Gly Lys Ser Thr Leu Ile Asn Leu Leu Ala Gly Leu Asp Thr Pro Asp
      35      40      45
Ser Gly Ser Val Tyr Ala Glu Gly Val Thr Val Ser Asp Gln Ser Glu
      50      55      60
Ala Ser Arg Ala Gln Phe Arg Leu Arg His Ile Ala Val Ile Phe Gln
65      70      75      80
Asp Asp Asn Leu Ile Ala Glu Leu Thr Asn Thr Glu Asn Ile Ala Leu
      85      90      95
Pro Leu Trp Ala Gln Gly Thr Ser Lys Ser Asp Ala Thr Glu Ile Ala
      100      105      110
His Glu Ala Met Arg Lys Leu Gly Ile Glu Ser Leu Gly Arg Arg Tyr
      115      120      125
Pro Gly Glu Val Ser Gly Gly Gln Arg Gln Arg
      130      135

```

<210> 297

<211> 378

<212> DNA

<213> Homo sapiens

<400> 297

```

tacaccatcg gtgaccagat tgtcgaagct ctgcaggtgc actcgaagat gtccgacaag
60
gacgcttggg cgcgtgccat cgagctgctc gacttggtgg ggattccgaa tcccagggtg
120
cgtgccaaag cttttccgca cgagttttcc ggtggcatga ggcaacgagt cgtcatcgcc
180
atggccatcg cgaacgaccc tgacctcatc atcgccgacg agccgacgac ggccctcgac
240
gtgaccatcc aggcccagat tctcgatttg ctgcgcgtag cccagcgtga aaccatgcg
300
ggcgtcgtaa tgatcaccca cgacctcggt gtggtagctg gtctggctga cagggttgcc
360
gtgatgtatg ccggacgc
378

```

<210> 298

<211> 126

<212> PRT

<213> Homo sapiens

<400> 298

```

Tyr Thr Ile Gly Asp Gln Ile Val Glu Ala Leu Gln Val His Ser Lys
1      5      10      15
Met Ser Asp Lys Asp Ala Trp Ala Arg Ala Ile Glu Leu Leu Asp Leu
      20      25      30
Val Gly Ile Pro Asn Pro Glu Val Arg Ala Lys Ala Phe Pro His Glu
      35      40      45
Phe Ser Gly Gly Met Arg Gln Arg Val Val Ile Ala Met Ala Ile Ala
      50      55      60
Asn Asp Pro Asp Leu Ile Ala Asp Glu Pro Thr Thr Ala Leu Asp
65      70      75      80
Val Thr Ile Gln Ala Gln Ile Leu Asp Leu Arg Val Ala Gln Arg

```

```

      85              90              95
Glu Thr His Ala Gly Val Val Met Ile Thr His Asp Leu Gly Val Val
      100              105              110
Ala Gly Leu Ala Asp Arg Val Ala Val Met Tyr Ala Gly Arg
      115              120              125

```

<210> 299
 <211> 368
 <212> DNA
 <213> Homo sapiens

```

<400> 299
gtgcacgggtt tcgttggcat ggcgaatgac cgggagaact tgcgttttga tccgagactt
60
ccagcccaat ggacgtcgat caaacaccac atgtcattg gcgactctca catgctcggtt
120
ttcctggaac gtgacgcat tacgttccag attctgtcgg gccatgaccg cgacgtgaca
180
gtgcgcggtg agctctacca cattgggggtt gagccggtga gggtgccggt gtccgatcag
240
gggcccgttg gtcctagcct gcgcgttacc catccgatct cgggggttgcg tcgagctgac
300
ggttctctta tcaactgcaga agttcccggc agcattgctg agacgattgg gtcttctccg
360
atctcgac
368

```

<210> 300
 <211> 122
 <212> PRT
 <213> Homo sapiens

```

<400> 300
Val His Gly Phe Val Gly Met Arg Asn Asp Arg Glu Asn Leu Arg Phe
1      5      10      15
Asp Pro Arg Leu Pro Ala Gln Trp Thr Ser Ile Lys His His Met Leu
20      25      30
Ile Gly Asp Ser His Met Leu Val Phe Leu Glu Arg Asp Ala Ile Thr
35      40      45
Phe Gln Ile Leu Ser Gly His Asp Arg Asp Val Thr Val Arg Gly Glu
50      55      60
Leu Tyr His Ile Gly Val Glu Pro Val Arg Val Pro Leu Ser Asp Gln
65      70      75      80
Gly Pro Leu Arg Pro Ser Leu Arg Val Thr His Pro Ile Ser Gly Leu
85      90      95
Arg Arg Ala Asp Gly Ser Leu Ile Thr Ala Glu Val Pro Gly Ser Ile
100     105     110
Ala Glu Thr Ile Gly Ser Ser Pro Ile Ser
115     120

```

<210> 301
 <211> 456
 <212> DNA
 <213> Homo sapiens

<400> 301
 ggccgggtta ttgccgccc gtttgtcggg gaaaccggc agaccttga gcgcaccggc
 60
 aaccggcgcg actatttcgt accgcccgc gaaccgacct tgctcgacag gcttacggac
 120
 gcggggccga cggatgatgc aatcggcaag attggtgata tctacgcga caaaggcgtg
 180
 tctcaggtgc gtaaggcaat ggcaatattg gccttggttc atgaaacact cattgccatg
 240
 gacgacgcgc aggacggcga tctggtcttc accaacttcg tggatttcga catgctctac
 300
 gggcatcgca gggatgtgcc cggctatgcc gccgcgctcg aggccttcga ccggaggctg
 360
 ccggaagcca tggcgaaatt gcggacgggc gatcttctga tcttgacagc cgatcatggc
 420
 tgcgaccga ccctcaaggg aaccgaccac acgcgt
 456

<210> 302
 <211> 152
 <212> PRT
 <213> Homo sapiens

<400> 302
 Gly Arg Val Ile Ala Arg Pro Phe Val Gly Glu Thr Arg Gln Thr Phe
 1 5 10 15
 Glu Arg Thr Gly Asn Arg Arg Asp Tyr Ser Val Pro Pro Pro Glu Pro
 20 25 30
 Thr Leu Leu Asp Arg Leu Thr Asp Ala Gly Arg Thr Val Ile Ala Ile
 35 40 45
 Gly Lys Ile Gly Asp Ile Tyr Ala His Lys Gly Val Ser Gln Val Arg
 50 55 60
 Lys Ala Met Ala Ile Leu Ala Leu Phe Asp Glu Thr Leu Ile Ala Met
 65 70 75 80
 Asp Asp Ala Gln Asp Gly Asp Leu Val Phe Thr Asn Phe Val Asp Phe
 85 90 95
 Asp Met Leu Tyr Gly His Arg Arg Asp Val Pro Gly Tyr Ala Ala Ala
 100 105 110
 Leu Glu Ala Phe Asp Arg Arg Leu Pro Glu Ala Met Ala Lys Leu Arg
 115 120 125
 Thr Gly Asp Leu Leu Ile Leu Thr Ala Asp His Gly Cys Asp Pro Thr
 130 135 140
 Leu Lys Gly Thr Asp His Thr Arg
 145 150

<210> 303
 <211> 402
 <212> DNA
 <213> Homo sapiens

<400> 303
 nncgtgggca tcgaggagtt cctcgacatg aagtatcacg cgacgccgat tcatcgctgc
 60

tgacagcggt tttccggaac acatcagcgt tcagacagga gcgaggagac catgtacctg
 120
 ggtgctcagc tggtcagtga cagcgagtac gagcagcgcc tgagacgtgt ccgtgagctc
 180
 atggaccgtc aggggtctgtc ggcgatcatc gtcaccgatc cggccaacat cttctatctg
 240
 atcggttaca acgcctggtc gttctacacc ccgcagatgc tgttcgtgcc gatcgacgga
 300
 gagatgggtcc tctacgctcg cgagatggat cgcattggcg acatcngcac gacgtcgttg
 360
 cccgccgatc agatcgtcgg ttaccgggag agttatgtgc ac
 402

<210> 304

<211> 97

<212> PRT

<213> Homo sapiens

<400> 304

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Tyr | Leu | Gly | Ala | Gln | Leu | Phe | Ser | Asp | Ser | Glu | Tyr | Glu | Gln | Arg |
| 1 | | | | 5 | | | | | 10 | | | | | 15 | |
| Leu | Arg | Arg | Val | Arg | Glu | Leu | Met | Asp | Arg | Gln | Gly | Leu | Ser | Ala | Ile |
| | | | 20 | | | | | 25 | | | | | 30 | | |
| Ile | Val | Thr | Asp | Pro | Ala | Asn | Ile | Phe | Tyr | Leu | Ile | Gly | Tyr | Asn | Ala |
| | | 35 | | | | | 40 | | | | | 45 | | | |
| Trp | Ser | Phe | Tyr | Thr | Pro | Gln | Met | Leu | Phe | Val | Pro | Ile | Asp | Gly | Glu |
| | 50 | | | | | 55 | | | | | 60 | | | | |
| Met | Val | Leu | Tyr | Ala | Arg | Glu | Met | Asp | Arg | Met | Ala | His | Ile | Xaa | Thr |
| 65 | | | | | 70 | | | | 75 | | | | 80 | | |
| Thr | Ser | Leu | Pro | Ala | Asp | Gln | Ile | Val | Gly | Tyr | Pro | Glu | Ser | Tyr | Val |
| | | | 85 | | | | | 90 | | | | | 95 | | |

His

<210> 305

<211> 375

<212> DNA

<213> Homo sapiens

<400> 305

nnacgcgtcg gttccgcac gagcgaccgg atcgcatcga cgagcacgct gcaccagtgc
 60
 gtgctgctct ggcaaatatg ggcgatcagc cggtagctt cgggacgtc gctcacctcg
 120
 gccgccattt cggatgcgac acgcgcgcct gcgcgtcgg cctccagcaa ctcgtcgagc
 180
 gtcgccacca gcgcggcgcg atcttcatgc ggagtcagat cggcgcgggc gtcaggcccc
 240
 tcgccatgcy tcggaatcga catgcagcac cctcctgcc ggatcgatgg cgtaatacgt
 300
 gcgacgggtac acggcgcggtg ttgcacgaac gtgcaaata gcgcgtgcct cgtgccatat
 360
 acgtcacatc atatg
 375

<210> 306
 <211> 125
 <212> PRT
 <213> Homo sapiens

<400> 306
 Xaa Arg Val Gly Ser Ala Ser Ser Asp Arg Ile Ala Ser Thr Ser Thr
 1 5 10 15
 Leu His Gln Cys Val Ser Ser Trp Arg Ile Trp Ala Ile Ser Arg Tyr
 20 25 30
 Ser Ser Gly Ser Ser Leu Thr Ser Ala Ala Ile Ser Asp Ala Thr Arg
 35 40 45
 Ala Pro Ala Arg Ser Ala Ser Ser Asn Ser Ser Ser Val Ala Thr Ser
 50 55 60
 Ala Ala Arg Ser Ser Cys Gly Val Arg Ser Ala Arg Ala Ser Gly Pro
 65 70 75 80
 Ser Pro Cys Val Gly Ile Asp Met Gln His Pro Pro Ala Arg Ile Asp
 85 90 95
 Gly Val Ile Arg Ala Thr Val His Gly Ala Cys Cys Thr Asn Val Gln
 100 105 110
 Ile Ser Ala Cys Leu Val Pro Tyr Thr Ser His His Met
 115 120 125

<210> 307
 <211> 685
 <212> DNA
 <213> Homo sapiens

<400> 307
 actagttctg gccgctcccc tggggctttg ggtaacaatt gtcagcccca cccatcctag
 60
 ggtaggaag gctattctct ttggccactc tcatacctaag acctatttgg agaacctctg
 120
 gggtttgagt ctttttttca gcagaatgag gcttgatccc gcattatagc acctcgcaca
 180
 tttgatgtct cttcttctca cccactcacc ccaccctggg gggtggggca aaaaagtggc
 240
 tcaaagctgc gggtcagagt tccttgtaaa caaggctcct ccctcactgt cctcaccctg
 300
 ctccagcaga gggagcagcg gaaggaccac tctgctgcag ccatgcttgt ttctaaccga
 360
 gcagaactgg acataatggg aacagggtct gaagacaatc aatccagggc tgcagtgggt
 420
 gctgagtctg gggaagcctc cacctggagg ggcagctggg cagtggcagc tcccttgga
 480
 tggctcagcc tctggacatc accccaccca accagagccc tggctcttgc tggatgtcca
 540
 cagatgagtg cctgggattg gtctcagcca ctatgggggg gatgtgcagg gagaggtgat
 600
 gagggagtga gcaggactgt ctatgtgcct ctgtcctcat cctgaggctt gggctctgaa
 660
 ttggtgctgc agcactggca cgcgt
 685

<210> 308
 <211> 100
 <212> PRT
 <213> Homo sapiens

<400> 308
 Met Leu Val Ser Asn Pro Ala Glu Leu Asp Ile Met Gly Thr Gly Ser
 1 5 10 15
 Glu Asp Asn Gln Ser Arg Ala Ala Val Gly Ala Glu Ser Gly Glu Ala
 20 25 30
 Ser Thr Trp Arg Gly Ser Trp Ala Val Ala Ala Pro Leu Glu Trp Leu
 35 40 45
 Ser Leu Trp Thr Ser Pro His Pro Thr Arg Ala Leu Ala Leu Ala Gly
 50 55 60
 Cys Pro Gln Met Ser Ala Trp Asp Trp Ser Gln Pro Leu Trp Gly Gly
 65 70 75 80
 Cys Ala Gly Arg Gly Asp Glu Gly Val Ser Arg Thr Val Tyr Val Pro
 85 90 95
 Leu Ser Ser Ser
 100

<210> 309
 <211> 432
 <212> DNA
 <213> Homo sapiens

<400> 309
 caggctcgta ctattcgat ccctgtgcat atggctcgagg tcatcaataa gctggctcgc
 60
 gtccagcgtc agatgctcca ggacctaggt cgtgagccca ccccggaaga gcttgccaac
 120
 gaactcgata tgaccgcaga gaaggctcatt gaggtgcaga aatacggtcg cgagccgac
 180
 tcgctgcata ccccaactggg tgaggatggc gattctgagt tcggtgacct tattgaggat
 240
 tccgaggcca tcgtgccagc agacgccgtc aacttcaccc tggtgcagga gcagctgcat
 300
 gatgtcctcg ataccttgtc cgagcgagag gccggtgtcg tgctgatgag attcggcttg
 360
 accgacggac agcccaagac cctggatgag atcggaag tctacggtgt tactcgggag
 420
 cgcacccgac ag
 432

<210> 310
 <211> 144
 <212> PRT
 <213> Homo sapiens

<400> 310
 Gln Ala Arg Thr Ile Arg Ile Pro Val His Met Val Glu Val Ile Asn
 1 5 10 15
 Lys Leu Ala Arg Val Gln Arg Gln Met Leu Gln Asp Leu Gly Arg Glu

20 25 30
 Pro Thr Pro Glu Glu Leu Ala Asn Glu Leu Asp Met Thr Ala Glu Lys
 35 40 45
 Val Ile Glu Val Gln Lys Tyr Gly Arg Glu Pro Ile Ser Leu His Thr
 50 55 60
 Pro Leu Gly Glu Asp Gly Asp Ser Glu Phe Gly Asp Leu Ile Glu Asp
 65 70 75 80
 Ser Glu Ala Ile Val Pro Ala Asp Ala Val Asn Phe Thr Leu Leu Gln
 85 90 95
 Glu Gln Leu His Asp Val Leu Asp Thr Leu Ser Glu Arg Glu Ala Gly
 100 105 110
 Val Val Ser Met Arg Phe Gly Leu Thr Asp Gly Gln Pro Lys Thr Leu
 115 120 125
 Asp Glu Ile Gly Lys Val Tyr Gly Val Thr Arg Glu Arg Ile Arg Gln
 130 135 140

<210> 311
 <211> 358
 <212> DNA
 <213> Homo sapiens

<400> 311
 acgcgtatcg aaaatatccc tccattatt accgctcgcc ctgaactgat ggctcatgaa
 60
 ctgacgccag aatctcttga tgcgagcctg gagtgggccc atgtgggtgt cattggctct
 120
 ggactgggac aacaagcgtg gggcaaaaaa gcgctacaaa aggtcgagaa ttgtcgtaaa
 180
 ccgatgctgt gggatgccga cgcgcttaac cttctggcaa tcaatcctga taaacgtcac
 240
 aatcgcatcc tgaagccaca ccccggcgag gccgcggcgc tgcttagctg cagcgtcgca
 300
 gaaattgaaa acgatcgctt acttntctgc gcacgtctgg taaaacggta acccgagt
 358

<210> 312
 <211> 116
 <212> PRT
 <213> Homo sapiens

<400> 312
 Thr Arg Ile Glu Asn Ile Pro Pro Ile Ile Thr Ala Arg Pro Glu Leu
 1 5 10 15
 Met Ala His Glu Leu Thr Pro Glu Ser Leu Asp Ala Ser Leu Glu Trp
 20 25 30
 Ala Asp Val Val Val Ile Gly Pro Gly Leu Gly Gln Gln Ala Trp Gly
 35 40 45
 Lys Lys Ala Leu Gln Lys Val Glu Asn Cys Arg Lys Pro Met Leu Trp
 50 55 60
 Asp Ala Asp Ala Leu Asn Leu Leu Ala Ile Asn Pro Asp Lys Arg His
 65 70 75 80
 Asn Arg Ile Leu Thr Pro His Pro Gly Glu Ala Ala Arg Leu Leu Ser
 85 90 95
 Cys Ser Val Ala Glu Ile Glu Asn Asp Arg Leu Leu Xaa Cys Ala Arg

100 105 110
 Leu Val Lys Arg
 115
 <210> 313
 <211> 347
 <212> DNA
 <213> Homo sapiens
 <400> 313
 ncaactgaaa gcattgagat gagcgacgtg ctgtccccct tccacccac caaggccaac
 60
 acccctggtg gcgaaccgcg caccatccgc acctcgaacg cgcacatcat tgccgtcacc
 120
 agtggcгааг gcggcgtggg caagaccttt gtctccgcca acctggccgc cgcgctgacc
 180
 cgcctgggac tgcgctgtct ggtactggac gccgacctgg gcctggccaa cttggacgtg
 240
 gtgctgaacc tctaccccaa ggtgacgtg cagcatgtgt tcaccggcaa ggctcgtg
 300
 caagacgcgg tggtcacggc ccccggcggc ttccatgtgc tgctagc
 347

<210> 314
 <211> 115
 <212> PRT
 <213> Homo sapiens

<400> 314
 Xaa Thr Glu Ser Ile Glu Met Ser Asp Val Leu Ser Pro Phe His Pro
 1 5 10 15
 Thr Lys Ala Asn Thr Pro Gly Gly Glu Pro Arg Thr Ile Arg Thr Ser
 20 25 30
 Asn Ala His Ile Ile Ala Val Thr Ser Gly Lys Gly Gly Val Gly Lys
 35 40 45
 Thr Phe Val Ser Ala Asn Leu Ala Ala Leu Thr Arg Leu Gly Leu
 50 55 60
 Arg Val Leu Val Leu Asp Ala Asp Leu Gly Leu Ala Asn Leu Asp Val
 65 70 75 80
 Val Leu Asn Leu Tyr Pro Lys Val Thr Leu His Asp Val Phe Thr Gly
 85 90 95
 Lys Ala Ser Leu Gln Asp Ala Val Val Thr Ala Pro Gly Gly Phe His
 100 105 110
 Val Leu Leu
 115

<210> 315
 <211> 544
 <212> DNA
 <213> Homo sapiens

<400> 315
 nnacgcgttc gtcaacagga aaacaacaac ggcttctcgc tggagggaac catgcttgcc
 60

gaagatatct acgcgatcat gctgttttca tcgctcatcc tggctgtccc ggggccatcc
 120
 aacaccttgc tgctcagcgc ccgtttccat ttccggctcgc tgcgggcggc gcccttcac
 180
 ctgcttgagg cgttgggcta ctcgctatcc atttcggcat ggggctgggt attggcgcgc
 240
 ctgtccgaga gcaatccatg gatcatcagt ctgaccaagg cactctgcgc gctatatgtg
 300
 gcgcttctgg cggtgaagac ctggaatgcc ntcgatccgc agtgcggggc cggttaacttc
 360
 cgccatgggc ccctgcccct gttcgtggca accctgtcga acccgaaggc gctgatcttc
 420
 gccagcgtga tctttcccgg caaggcgttc ctcgacttct ggaacaacta cacgatctcg
 480
 ctgctggcct tcctggttgt gctggcgccc atcgggatgc tttgggtcgg gctgggggcc
 540
 ggta
 544

<210> 316
 <211> 159
 <212> PRT
 <213> Homo sapiens

<400> 316
 Ile Tyr Ala Ile Met Leu Phe Ser Ser Leu Ile Leu Val Val Pro Gly
 1 5 10 15
 Pro Ser Asn Thr Leu Leu Leu Ser Ala Arg Phe His Phe Gly Ser Leu
 20 25 30
 Arg Ala Ala Pro Phe Ile Leu Leu Glu Ala Leu Gly Tyr Ser Leu Ser
 35 40 45
 Ile Ser Ala Trp Gly Trp Val Leu Ala Arg Leu Ser Glu Ser Asn Pro
 50 55 60
 Trp Ile Ile Ser Leu Thr Lys Ala Leu Cys Ala Leu Tyr Val Ala Leu
 65 70 75 80
 Leu Ala Val Lys Thr Trp Asn Ala Xaa Asp Pro Gln Cys Gly Ala Gly
 85 90 95
 Asn Phe Arg His Gly Pro Leu Pro Leu Phe Val Ala Thr Leu Ser Asn
 100 105 110
 Pro Lys Ala Leu Ile Phe Ala Ser Val Ile Phe Pro Gly Lys Ala Phe
 115 120 125
 Leu Asp Phe Trp Asn Asn Tyr Thr Ile Ser Leu Leu Ala Phe Leu Val
 130 135 140
 Val Leu Ala Pro Ile Gly Met Leu Trp Val Gly Leu Gly Ala Gly
 145 150 155

<210> 317
 <211> 343
 <212> DNA
 <213> Homo sapiens

<400> 317
 nggtcagcct ctcgcccagg caattctctt aagatacatg agctgctatg agtaccaaag
 60

ccagaggttt gtccactgag agaagcacat tggaaaagggg ggcgtggggc tgggactgtg
 120
 tggcacttta tgcacggggg gggcctaagg gggnggtcc accaaccatg cactgngggg
 180
 ggggtgtggg taacatgccg tgcattttgg ggggtgtgcca tgagtggcac accatggggg
 240
 tggcatgtgg ggcattgtat catgtggtgt tggcgagca aactcagctc ttacctggct
 300
 ggggccagcc tctaaaactt ctcacattgg gctcccttct gac
 343

<210> 318

<211> 98

<212> PRT

<213> Homo sapiens

<400> 318

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Ser | Thr | Lys | Ala | Arg | Gly | Leu | Ser | Thr | Glu | Arg | Ser | Thr | Leu | Glu |
| 1 | | | | 5 | | | | 10 | | | | | | 15 | |
| Arg | Gly | Ala | Trp | Ala | Trp | Asp | Cys | Val | Ala | Leu | Tyr | Ala | Arg | Gly | Gly |
| | | | 20 | | | | 25 | | | | | | 30 | | |
| Pro | Lys | Gly | Gly | Gly | Pro | Pro | Thr | Met | His | Xaa | Gly | Trp | Gly | Val | Gly |
| | | | 35 | | | | 40 | | | | | 45 | | | |
| Asn | Met | Pro | Cys | Ile | Leu | Gly | Val | Cys | His | Glu | Trp | His | Thr | Met | Gly |
| | | | 50 | | | | 55 | | | | 60 | | | | |
| Val | Ala | Cys | Gly | Ala | Cys | Met | His | Val | Val | Leu | Ala | Gln | Gln | Thr | Gln |
| 65 | | | | | 70 | | | | | 75 | | | | 80 | |
| Leu | Leu | Pro | Gly | Trp | Gly | Gln | Pro | Leu | Lys | Leu | Leu | Thr | Leu | Gly | Ser |
| | | | | 85 | | | | 90 | | | | | | 95 | |
| Leu | Leu | | | | | | | | | | | | | | |

<210> 319

<211> 429

<212> DNA

<213> Homo sapiens

<400> 319

gaattctcga tgtaccccct cccggcagtc ctattctcga gctgagcggg cacagtggcc
 60
 ccgttaacag tgtggcttgg ggtccacca gccagagcac gttgcgaaat ggacctagta
 120
 agggcatgat atgtacagga ggcgacgatg ctcagtgcct cgtatatgat ctgactagct
 180
 caactcttcg aacagcatct gctcaaggac ggcgctctcg aaacagtcca tataaataaa
 240
 gccattcacc gggaatagac ggatggcgtg tcggcgagca agtgccggtg ctcgcttata
 300
 cggccccgtc tatggtcaac aatgctagct ggctcggcat gcctgcgcca tcaaaacgca
 360
 catcgctaca gagcaaacac cgcagccttt accgcagctt actcagttag tggacttagt
 420
 atacgtccn
 429

<210> 320
 <211> 101
 <212> PRT
 <213> Homo sapiens

<400> 320
 Met Ile Cys Thr Gly Gly Asp Asp Ala Gln Cys Leu Val Tyr Asp Leu
 1 5 10 15
 Thr Ser Ser Thr Leu Arg Thr Ala Ser Ala Gln Gly Arg Arg Ser Arg
 20 25 30
 Asn Ser Pro Tyr Lys Gln Ser His Ser Pro Gly Ile Asp Gly Trp Arg
 35 40 45
 Val Gly Ala Glu Val Pro Val Leu Ala Tyr Thr Ala Pro Ser Met Val
 50 55 60
 Asn Asn Ala Ser Trp Leu Gly Met Pro Ala Pro Ser Lys Arg Thr Ser
 65 70 75 80
 Leu Gln Ser Lys His Arg Ser Leu Tyr Arg Ser Leu Leu Ser Glu Trp
 85 90 95
 Thr Glu Tyr Thr Ser
 100

<210> 321
 <211> 530
 <212> DNA
 <213> Homo sapiens

<400> 321
 ngtgcacgac gtgctcgcca agtccctcgg gtcctctaata gcatcaacg tggttcacgc
 60
 caccgtcgat gcgttgacgc agctcgagga gcccgaaagag gtcgcccgtc gccgcggcaa
 120
 gtccgttgag gagatcgccc cagcagccat gctgcgtgag cgcaaggagg ccgacgaggc
 180
 cgccgctgct gcccgcatgg aggaaaaggc ggggggtaac tgatgagcaa gctgaagatc
 240
 acccagatca agtctggcat cgctaccaag ccaaatacatc gtgagaccct gcgcagcctc
 300
 ggactgaagc gtattggtga caccgtcatc aaggaggacc gcccgaggtt ccgcggcatg
 360
 gtcgggaccg ttcgtcacct cgtcaccatg gaagaggtgg actgacatgg ctattgagct
 420
 ccatgacctc aagccccgtc ctggtgcccc caaggccaag acccggttg gtcgtggtga
 480
 gggttccaag ggtaagaccg ctggtcgagg taccaagggc accggtgcac
 530

<210> 322
 <211> 60
 <212> PRT
 <213> Homo sapiens

<400> 322
 Met Ser Lys Leu Lys Ile Thr Gln Ile Lys Ser Gly Ile Ala Thr Lys

```

      1           5           10           15
Pro Asn His Arg Glu Thr Leu Arg Ser Leu Gly Leu Lys Arg Ile Gly
      20           25           30
Asp Thr Val Ile Lys Glu Asp Arg Pro Glu Phe Arg Gly Met Val Arg
      35           40           45
Thr Val Arg His Leu Val Thr Met Glu Glu Val Asp
      50           55           60

```

<210> 323

<211> 468

<212> DNA

<213> Homo sapiens

<400> 323

```

ntccggaccc gctgtggcca cgtattctgc cggttcctgta ttgctaccag tctaaagaac
60
aacaagtgga cctgtcctta ttgccgggca tatcttcctt cagaaggagt tccagcaact
120
gatgtagcca aaagaatgaa atcagagtat aagaactgcg ctgagtgtga caccctgggt
180
tgccctcagt aaatgagggc acatattcgg acttgtcaga agtacataga taagtatgga
240
ccactacaag aacttgagga gacagcagca aggtgtgtat gtcccttttg tcagagggaa
300
ctgtatgaag acagcttgct ggatcattgt attactcatc acagatcgga acggaggcct
360
gtgttctgtc cactttgcca ttttaataccc gatgagaatc caagcagctt cagtggcagt
420
ttaataagac atctgcaagt tagtcacact ttggtttatg atgatttc
468

```

<210> 324

<211> 156

<212> PRT

<213> Homo sapiens

<400> 324

```

Xaa Arg Thr Arg Cys Gly His Val Phe Cys Arg Ser Cys Ile Ala Thr
  1           5           10           15
Ser Leu Lys Asn Asn Lys Trp Thr Cys Pro Tyr Cys Arg Ala Tyr Leu
      20           25           30
Pro Ser Glu Gly Val Pro Ala Thr Asp Val Ala Lys Arg Met Lys Ser
      35           40           45
Glu Tyr Lys Asn Cys Ala Glu Cys Asp Thr Leu Val Cys Leu Ser Glu
      50           55           60
Met Arg Ala His Ile Arg Thr Cys Gln Lys Tyr Ile Asp Lys Tyr Gly
      65           70           75           80
Pro Leu Gln Glu Leu Glu Glu Thr Ala Ala Arg Cys Val Cys Pro Phe
      85           90           95
Cys Gln Arg Glu Leu Tyr Glu Asp Ser Leu Leu Asp His Cys Ile Thr
      100          105          110
His His Arg Ser Glu Arg Arg Pro Val Phe Cys Pro Leu Cys His Leu
      115          120          125
Ile Pro Asp Glu Asn Pro Ser Ser Phe Ser Gly Ser Leu Ile Arg His

```

130 135 140
 Leu Gln Val Ser His Thr Leu Val Tyr Asp Asp Phe
 145 150 155

<210> 325
 <211> 374
 <212> DNA
 <213> Homo sapiens

<400> 325
 acgcgtgaag ggaggacgag gaagtaacgg gaagcacaag gccgctgctg gggagatggc
 60
 actggagccc cctaggaagc atctcacagg ctgtggccct tggcacgggg atctggggcc
 120
 aggtcgagcg caggtctggg tatcatgcga gtgcgggctc gctggggcgg gaaagagttt
 180
 ggagctctgc tcccagggaa tccccactcc cgcagatgac ttgcccgaga gagttctgct
 240
 ggtggatttt gatggaaatt ctatttgatc gcacccactt ggttcactgt gtgcttccgg
 300
 gtccccaggt tttaggtgct tcatgccttg ctgggaacga gacacgctcc tgccctcagt
 360
 gaatcttcag tcta
 374

<210> 326
 <211> 108
 <212> PRT
 <213> Homo sapiens

<400> 326
 Met Lys His Leu Lys Pro Gly Asp Pro Glu Ala His Ser Glu Pro Ser
 1 5 10 15
 Gly Cys Asp Gln Ile Glu Phe Pro Ser Lys Ser Thr Ser Arg Thr Leu
 20 25 30
 Ser Gly Lys Ser Ser Ala Gly Val Gly Ile Pro Trp Glu Gln Ser Ser
 35 40 45
 Lys Leu Phe Pro Ala Pro Ala Ser Pro His Ser His Asp Thr Gln Thr
 50 55 60
 Cys Ala Arg Pro Gly Pro Arg Ser Pro Cys Gln Gly Pro Gln Pro Val
 65 70 75 80
 Arg Cys Phe Leu Gly Gly Ser Ser Ala Ile Ser Pro Ala Ala Ala Leu
 85 90 95
 Cys Phe Pro Leu Leu Pro Arg Pro Pro Phe Thr Arg
 100 105

<210> 327
 <211> 538
 <212> DNA
 <213> Homo sapiens

<400> 327
 cactataaaa tccagtttgg ggcccgtgtt ctttcctatt ggtctgtcag gtgaaaaact
 60

ccggctgggg gaaaagcgtc cggtggtttg ttggtaaaga gggcgctga tgggctctgg
 120
 ggaatggagg atggcgcacc ggctgtgggt ggactgtgga aacggggggg ggcagtgccg
 180
 gggtagttgt cctgctggtc tggttttggg atcctgggct ggagaaatgc gatccaaaag
 240
 agctcgggat gggctcagag cgaccacga aaataccagg ggccaagtaa aatgaacca
 300
 ccctttaaca gtgcacaaag cgctggcaca cggtcacgt ctggtgacgc aggctgcccg
 360
 aagcgctcca accattttgc aaacctggga gagcaagagg ggctctgcag gtctagccgc
 420
 cgccccctgc cactctggc cagccggagt tttcaccta cagaccaata ggaaagaaca
 480
 cgggccccaa actggatttt atagtctgag ctctcagcat ctaaggaatg atatgcc
 538

<210> 328

<211> 125

<212> PRT

<213> Homo sapiens

<400> 328

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Val | Gly | Ala | Leu | Arg | Ala | Ala | Cys | Val | Thr | Arg | Arg | Gly | Pro | Cys |
| 1 | | | | 5 | | | | | 10 | | | | | 15 | |
| Ala | Ser | Ala | Leu | Cys | Thr | Val | Lys | Gly | Trp | Val | His | Phe | Thr | Trp | Pro |
| | | | 20 | | | | | 25 | | | | | 30 | | |
| Leu | Val | Phe | Ser | Trp | Val | Ala | Leu | Ser | Pro | Ser | Arg | Ala | Leu | Leu | Asp |
| | 35 | | | | | | 40 | | | | | 45 | | | |
| Arg | Ile | Ser | Pro | Ala | Gln | Asp | Pro | Lys | Thr | Arg | Pro | Ala | Gly | Gln | Leu |
| | 50 | | | | 55 | | | | | 60 | | | | | |
| Pro | Arg | His | Cys | His | Pro | Pro | Phe | Pro | Gln | Ser | Thr | His | Ser | Arg | Cys |
| 65 | | | | 70 | | | | | 75 | | | | | 80 | |
| Ala | Ile | Leu | His | Ser | Pro | Glu | Pro | Ile | Thr | His | Pro | Leu | Tyr | Gln | Gln |
| | | | 85 | | | | | | 90 | | | | | 95 | |
| Thr | Thr | Gly | Arg | Phe | Ser | Pro | Ser | Arg | Ser | Phe | Ser | Pro | Asp | Arg | Pro |
| | | 100 | | | | | | 105 | | | | | 110 | | |
| Ile | Gly | Lys | Asn | Thr | Gly | Pro | Lys | Leu | Asp | Phe | Ile | Val | | | |
| | | 115 | | | | | 120 | | | | | 125 | | | |

<210> 329

<211> 407

<212> DNA

<213> Homo sapiens

<400> 329

tccggagagt tccctcccca ggaattcctt ctaagaatcc atgtggaaat agagcctgaa
 60
 gctcttcagt ctttctgctc cactgagcag tgttttcttg atacccttgg tatectgcc
 120
 gcagcctcgt tatgactcct aactccattg ccctccatgg cccctgggag ctctctctct
 180
 ctttctctcc aggtagtaga gactgcttc tggcttcttg tgcacagaag ggtttccac
 240

agctgagagc tgggctccta ctgacatagt tatttccttt atatacctgcc ccaccttctt
 300
 ctggtagcac acagcaacct tgcatagttag ctggtatcat taccttccca atcaacaggc
 360
 cttgatttct tataggactt tttctctcag atttacattg cttcttt
 407

<210> 330
 <211> 113
 <212> PRT
 <213> Homo sapiens

<400> 330
 Met Ile Pro Ala Thr Met Gln Gly Cys Cys Val Leu Pro Glu Glu Gly
 1 5 10 15
 Gly Ala Gly Tyr Lys Gly Asn Asn Tyr Val Ser Arg Ser Pro Ala Leu
 20 25 30
 Ser Cys Gly Lys Pro Phe Cys Ala Gln Glu Ala Arg Ser Ser Ala Leu
 35 40 45
 Leu Pro Gly Glu Lys Glu Arg Glu Ser Ala Gln Gly Pro Trp Arg Ala
 50 55 60
 Met Glu Leu Gly Val Ile Thr Arg Leu Leu Ala Gly Tyr Gln Gly Tyr
 65 70 75 80
 Gln Glu Asn Thr Ala Gln Trp Ser Arg Lys Thr Glu Glu Leu Gln Ala
 85 90 95
 Leu Phe Pro His Gly Phe Leu Glu Gly Ile Pro Gly Glu Gly Thr Leu
 100 105 110
 Arg

<210> 331
 <211> 523
 <212> DNA
 <213> Homo sapiens

<400> 331
 tgtaccgaac ctgctggtct cgagggcctt gctgggctcg tcgtacgcac agctgacgaa
 60
 tccaccggcc cccatcccg cgccactttc gctgaggcca tggagtcgat cggagccagc
 120
 tacgacggat cggccgggtt ggccggaagt cagctcggcg tcgatgtgcc cgtgacaagg
 180
 ttgcagcgag cggctgaact cttcgtcgaa ttgttgaaca ccacgagcct ggttgaagag
 240
 gacatcgccc gtcagatcga cgcggcgcca gcctccctgg cccagaccag ccagcgcgga
 300
 tcggccctag ccgagatggc agcagcacgt gcgctatggc cagtggggtc acggctcgtcc
 360
 ctgcccacga tcggtaccct ctgctcgggtg gaaaagctca acgccgcagc cgcacgagaa
 420
 ttctgggccg cgcactggac gatctccgat gccgtgctgg tggttgccgg agagggagtc
 480
 gaggacctcg acttgtcaat attcaaggag tggacgacca gct
 523

<210> 332
 <211> 174
 <212> PRT
 <213> Homo sapiens

<400> 332
 Cys Thr Glu Pro Ala Gly Leu Glu Gly Leu Ala Gly Leu Val Val Arg
 1 5 10 15
 Thr Ala Asp Glu Ser Thr Gly Pro His Pro Gly Ala Thr Phe Ala Glu
 20 25 30
 Ala Met Glu Ser Ile Gly Ala Ser Tyr Asp Gly Ser Ala Gly Leu Ala
 35 40 45
 Gly Ser His Val Gly Val Asp Val Pro Val Thr Arg Phe Asp Ala Ala
 50 55 60
 Ala Glu Leu Phe Val Glu Leu Leu Asn Thr Thr Ser Leu Val Glu Glu
 65 70 75 80
 Asp Ile Ala Arg Gln Ile Asp Ala Ala Arg Ala Ser Leu Ala Gln Thr
 85 90 95
 Ser Gln Arg Gly Ser Ala Leu Ala Glu Met Ala Ala Ala Arg Ala Leu
 100 105 110
 Trp Pro Val Gly Ser Arg Ser Ser Leu Pro Thr Ile Gly Thr Leu Ser
 115 120 125
 Ser Val Glu Lys Leu Asn Ala Ala Ala Arg Glu Phe Trp Ala Ala
 130 135 140
 His Trp Thr Ile Ser Asp Ala Val Leu Val Val Ala Gly Glu Gly Val
 145 150 155 160
 Glu Asp Leu Asp Leu Ser Ile Phe Lys Glu Trp Thr Thr Ser
 165 170

<210> 333
 <211> 372
 <212> DNA
 <213> Homo sapiens

<400> 333
 nntgttcgtc gtgtcgaccc ggaactcaag gcccgagcga tgacgggtgaa ggtgccaacc
 60
 gatccccatc accgcccggg agttccattg aagtctgcga aggaccgtat ggacatcatt
 120
 tctgcttacc gagaactcgg aagctatcgc gccgcagccg aggtgtgcgg caccacccac
 180
 aagaccgtca agcgggtggt cgatcggttt gaagccggcg atccacccac cgggtggcaag
 240
 gaacggggccc gcaactacga tgcgggtggcc cagctcgtcg cgcagcgagt cgcgcgggtca
 300
 cacggccgga tcaactgcaa acggctgcta ccggtagcgc gagcggcagg atatgagggg
 360
 tcggcgcgga at
 372

<210> 334
 <211> 88
 <212> PRT

<213> Homo sapiens

<400> 334

```

Met Asp Ile Ile Ser Ala Tyr Arg Glu Leu Gly Ser Tyr Arg Ala Ala
 1           5           10           15
Ala Glu Val Cys Gly Thr Thr His Lys Thr Val Lys Arg Val Val Asp
      20           25           30
Arg Phe Glu Ala Gly Asp Pro Pro Thr Gly Gly Lys Glu Arg Ala Arg
      35           40           45
Asn Tyr Asp Ala Val Ala Gln Leu Val Ala Gln Arg Val Ala Arg Ser
      50           55           60
His Gly Arg Ile Thr Ala Lys Arg Leu Leu Pro Val Ala Arg Ala Ala
      65           70           75           80
Gly Tyr Glu Gly Ser Ala Arg Asn
                        85

```

<210> 335

<211> 356

<212> DNA

<213> Homo sapiens

<400> 335

```

gtgcacgcct tgctgggcga gggcgatgcg cctgcgcgca ccttcgtgga cggtagcttt
60
ggcaggggag ggcattcgcg gctcatcctg cagcgggttg ggccgcaagg ccgcctgggtg
120
gcgttcgaca aggacaccga agccattcaa gcagcggcgc gcatcacgga tgcgcgcttt
180
tccatcnggc accagggggt cagccatctc ggggaactgc ccgccgccag cgtgtccggt
240
gtgctgctgg acctgggcgt gagctccccg cagatcgacg acccccagcg cgggttcagt
300
tttcgtttcg atgggtccgct ggacatgcgc atggacacca ctccgatgca tggatg
356

```

<210> 336

<211> 118

<212> PRT

<213> Homo sapiens

<400> 336

```

Val His Ala Leu Leu Gly Glu Gly Asp Ala Pro Ala Arg Thr Phe Val
 1           5           10           15
Asp Gly Thr Phe Gly Arg Gly Gly His Ser Arg Leu Ile Leu Gln Arg
      20           25           30
Leu Gly Pro Gln Gly Arg Leu Val Ala Phe Asp Lys Asp Thr Glu Ala
      35           40           45
Ile Gln Ala Ala Ala Arg Ile Thr Asp Ala Arg Phe Ser Ile Xaa His
      50           55           60
Gln Gly Phe Ser His Leu Gly Glu Leu Pro Ala Ala Ser Val Ser Gly
      65           70           75           80
Val Leu Leu Asp Leu Gly Val Ser Ser Pro Gln Ile Asp Asp Pro Gln
      85           90           95
Arg Gly Phe Ser Phe Arg Phe Asp Gly Pro Leu Asp Met Arg Met Asp

```

100
Thr Thr Pro Met His Gly
115

105

110

<210> 337
<211> 447
<212> DNA
<213> Homo sapiens

<400> 337
cagcctctct ccgaccgcgc cgggtgtgaag cacgggcatg ccggtgtgca agtggcacca
60
cagccaaaac agcgagctca cacttcaaac tccttcaaag accccaggcc tctgtaagaa
120
ccgctcatct ctgtgcccac agtcccccg cttccatgtg acccagaaat ggaaccacgc
180
agcagaggcg gggatcacag gtgaagcagc tgtgaacatt tgcttcaggc ttctgtgcaa
240
acaggcgcca tcatgtcagc cggtagcagc gagcaacgtg cgtgggtcag ggggtggcca
300
cacgtccaac ttataagaa atgacagatt ccctgatggc catagggatc tgcagggcca
360
gcagcaggca taggacttcc ggtggccctg cgtcttcac cacttgagt attgtcaggg
420
tttctgtact gtttttacag ccaattg
447

<210> 338
<211> 111
<212> PRT
<213> Homo sapiens

<400> 338
Met Pro Val Cys Lys Trp His His Ser Gln Asn Ser Glu Leu Thr Leu
1 5 10 15
Gln Thr Pro Ser Lys Thr Pro Gly Leu Cys Lys Asn Arg Ser Ser Leu
20 25 30
Cys Pro Gln Leu Pro Arg Phe His Val Thr Gln Lys Trp Asn His Ala
35 40 45
Ala Glu Ala Gly Ile Thr Gly Glu Ala Ala Val Asn Ile Cys Phe Arg
50 55 60
Leu Leu Cys Lys Gln Ala Pro Ser Cys Gln Pro Val Ser Arg Ser Asn
65 70 75 80
Val Arg Gly Ser Gly Gly Gly His Thr Ser Asn Phe Ile Arg Asn Asp
85 90 95
Arg Phe Pro Asp Gly His Arg Asp Leu Gln Gly Gln Gln Ala
100 105 110

<210> 339
<211> 588
<212> DNA
<213> Homo sapiens

<400> 339

tctagaatga agcgctgtat cctagcaccg gcagacgtac caagactatc aaggcgctca
 60
 gatcgtttat cctgcagttg ccattcatca gacaaatcca gtggaacca atggaagaca
 120
 ccgacctgca agcgctgatg gccagactcg aattgctaata tgatcgggtc gagcaactta
 180
 agagtcaaaa cggactccta ttagctcagg aaaagacctg ggcgcganaa cgcgtcacc
 240
 tcattgaaaa aaacgaaatc gcccggcgta aggtcgaatc gatgatttcg cgcctgaagg
 300
 ccctggagca agactatgag ttaagcaata gcgttacgtg cagatcctcg acaaagaata
 360
 ttgatcatc tgccccagg aagaacgcag cacctggtga gtgctgcccg ctacctggaa
 420
 ggccaaaagg cgtgaaatcc gcagcagcgg caaagtcatc ggtgccgacc gcatcgccgt
 480
 gatggccgcg ctgaacatca cccacgatct gctgcataag caggaacggc ctgacgttca
 540
 ggccagcggc tcaacgcgcg agcaagtgcg tgacctgctg gaacgcgt
 588

<210> 340
 <211> 123
 <212> PRT
 <213> Homo sapiens

<400> 340
 Met Glu Asp Thr Asp Leu Gln Ala Leu Met Ala Arg Leu Glu Leu Leu
 1 5 10 15
 Ile Asp Arg Val Glu Gln Leu Lys Ser Gln Asn Gly Leu Leu Leu Ala
 20 25 30
 Gln Glu Lys Thr Trp Ala Arg Xaa Arg Ala His Leu Ile Glu Lys Asn
 35 40 45
 Glu Ile Ala Arg Arg Lys Val Glu Ser Met Ile Ser Arg Leu Lys Ala
 50 55 60
 Leu Glu Gln Asp Tyr Glu Leu Ser Asn Ser Val Thr Cys Arg Ser Ser
 65 70 75 80
 Thr Lys Asn Ile Arg Ser Ser Ala Pro Arg Lys Asn Ala Ala Pro Gly
 85 90 95
 Glu Cys Cys Pro Leu Pro Gly Arg Pro Lys Gly Val Lys Ser Ala Ala
 100 105 110
 Ala Ala Lys Ser Ser Val Pro Thr Ala Ser Pro
 115 120

<210> 341
 <211> 401
 <212> DNA
 <213> Homo sapiens

<400> 341
 ngccgcgcgg cctacctgct gtacctggcc tatgccacct ggcgtgaccg ctcggccttt
 60
 gcaatgaacg acacgccgac agttgcgacc gcgcgcagcc tgatcctgcg tggcttcttg
 120

ctgaacattc ttaaccccaa gctgacaatt ttcttctctgg ccttcctgcc tcaattcgta
 180
 acgccaggcg gcaccgcgcc ggccttgacg atgctggtac tgagcggcgt gttcatggcg
 240
 atgacgcttg cagtgtttgt gctgtatggc ctgttggcga atgtgtttcg tcgtgcagtg
 300
 gtcgagtcgc cacgtgtgca gaactggctg cgacgcagtt ttgccacggc ctttgccggg
 360
 ctgggggtga acctggcggt tgcgcgagcg tgaggacgcg t
 401

<210> 342

<211> 130

<212> PRT

<213> Homo sapiens

<400> 342

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Xaa | Arg | Ala | Ala | Tyr | Leu | Leu | Tyr | Leu | Ala | Tyr | Ala | Thr | Trp | Arg | Asp |
| 1 | | | | 5 | | | | | 10 | | | | | 15 | |
| Arg | Ser | Ala | Phe | Ala | Met | Asn | Asp | Thr | Pro | Thr | Val | Ala | Thr | Ala | Arg |
| | | | 20 | | | | | 25 | | | | | 30 | | |
| Ser | Leu | Ile | Leu | Arg | Gly | Phe | Leu | Leu | Asn | Ile | Leu | Asn | Pro | Lys | Leu |
| | | 35 | | | | | 40 | | | | | 45 | | | |
| Thr | Ile | Phe | Phe | Leu | Ala | Phe | Leu | Pro | Gln | Phe | Val | Thr | Pro | Gly | Gly |
| | 50 | | | | | 55 | | | | | 60 | | | | |
| Thr | Ala | Pro | Ala | Leu | Gln | Met | Leu | Val | Leu | Ser | Gly | Val | Phe | Met | Ala |
| 65 | | | | | 70 | | | | | 75 | | | | 80 | |
| Met | Thr | Leu | Ala | Val | Phe | Val | Leu | Tyr | Gly | Leu | Leu | Ala | Asn | Val | Phe |
| | | | 85 | | | | | | 90 | | | | 95 | | |
| Arg | Arg | Ala | Val | Val | Glu | Ser | Pro | Arg | Val | Gln | Asn | Trp | Leu | Arg | Arg |
| | | | 100 | | | | | 105 | | | | | 110 | | |
| Ser | Phe | Ala | Thr | Ala | Phe | Ala | Gly | Leu | Gly | Leu | Asn | Leu | Ala | Phe | Ala |
| | | 115 | | | | | 120 | | | | | 125 | | | |
| Gln | Arg | | | | | | | | | | | | | | |
| | 130 | | | | | | | | | | | | | | |

<210> 343

<211> 389

<212> DNA

<213> Homo sapiens

<400> 343

gtgttgcgca actacatggc gtccctgccg ttcagcgtgg tcgagtcggc gcgcacgcac
 60
 ggggtgctcca acttcagat cttctggaag ctgatcgccc cgatggcgat gccggcgatg
 120
 gcggcggttcg cgaccctgca gttcctgtgg gtgtggaacg acctgctcat cgccaagctc
 180
 ttccctacca acgacaaccc cacggtgatc gtcaagctcc aacagcttcc cnngggcccc
 240
 aaggcccagg gtgcggagct gctgacggcg ggcgccttca tctccatcgt gctacccatg
 300
 atcgtcttct tcgtgctcca gaacttctg gtgcgcggta tgacgtcggg tgccgtcaag
 360

gggtgaccgc tcaactgcag tggcccggg
389

<210> 344
<211> 121
<212> PRT
<213> Homo sapiens

<400> 344
Val Leu Arg Asn Tyr Met Ala Ser Leu Pro Phe Ser Val Val Glu Ser
1 5 10 15
Ala Arg Ile Asp Gly Cys Ser Asn Phe Gln Ile Phe Trp Lys Leu Ile
20 25 30
Ala Pro Met Ala Met Pro Ala Met Ala Ala Phe Ala Thr Leu Gln Phe
35 40 45
Leu Trp Val Trp Asn Asp Leu Leu Ile Ala Lys Leu Phe Leu Thr Asn
50 55 60
Asp Asn Pro Thr Val Ile Val Lys Leu Gln Gln Leu Ser Xaa Gly Pro
65 70 75 80
Lys Ala Gln Gly Ala Glu Leu Leu Thr Ala Gly Ala Phe Ile Ser Ile
85 90 95
Val Leu Pro Met Ile Val Phe Phe Val Leu Gln Asn Phe Leu Val Arg
100 105 110
Gly Met Thr Ser Gly Ala Val Lys Gly
115 120

<210> 345
<211> 360
<212> DNA
<213> Homo sapiens

<400> 345
ctagtacttt atgctgatgg tgaacgtcgt tacatccttg cccttaaagg catggttgct
60
ggtgatgtga tccaatctgg tgaagatgca tcaattaaag taggtaactg cttaccgatg
120
cgtaatatcc cagttggtac aacagtacac gctgtagaaa tgaaacctgc taaaggtgca
180
caaattgcac gttctgctgg ttcttacagc caaattatag ctgctgatgg tgcttacgtt
240
actctacgtt tacgtagtgg tgaaatgcgt aaaatccctg ctgagtgtcg tgcaacaatc
300
ggtgaagttg gtaatgcaga acatatgcta cgtcaactag gtaaagctgg tgctacgcgt
360

<210> 346
<211> 120
<212> PRT
<213> Homo sapiens

<400> 346
Leu Val Leu Tyr Ala Asp Gly Glu Arg Arg Tyr Ile Leu Ala Pro Lys
1 5 10 15
Gly Met Val Ala Gly Asp Val Ile Gln Ser Gly Glu Asp Ala Ser Ile

```

      20      25      30
Lys Val Gly Asn Cys Leu Pro Met Arg Asn Ile Pro Val Gly Thr Thr
      35      40      45
Val His Ala Val Glu Met Lys Pro Ala Lys Gly Ala Gln Ile Ala Arg
      50      55      60
Ser Ala Gly Ser Tyr Ser Gln Ile Ile Ala Arg Asp Gly Ala Tyr Val
65      70      75      80
Thr Leu Arg Leu Arg Ser Gly Glu Met Arg Lys Ile Pro Ala Glu Cys
      85      90      95
Arg Ala Thr Ile Gly Glu Val Gly Asn Ala Glu His Met Leu Arg Gln
      100      105      110
Leu Gly Lys Ala Gly Ala Thr Arg
      115      120

```

<210> 347

<211> 565

<212> DNA

<213> Homo sapiens

<400> 347

```

accggtgatg ccaaaggtgc tgtgacaagg ggattcatcg gttcgggcaa ggtcgtcacg
60
gcagctgccg tcatcatgat ttcggtgttc gtcttcttca tccccgaggg catgaacgcc
120
atcaaggaaa tcgccctggc cctggccgtc gggatcctca cggatgcctt cttggtgcgg
180
atgaccctcg tcccggccgt gatggccctg ctaggtgaca aggcattggtg gttgcccggg
240
tggctggatc gacgcctacc ccgcctcgac atcgaggag aagggatcac ccacgaggaa
300
aagctggccg cctggcccac agcggatcac accgaggccc tgcacgccga ggggatcggg
360
gtggaggggc tcttcgaagg cctcgatctg cacgtcgaac cgcgtcaggt gcaagccgtc
420
gtcggatcgc agaacagtgt ctgcccgtc ctgctggcga tccggggacg gctgcccttg
480
gatcacggcc ggatgaggtc gggaggattg ctgctaccgc agcgggcttc cagagtgcgt
540
cgggtgacgt gggttcctga cgcgt
565

```

<210> 348

<211> 188

<212> PRT

<213> Homo sapiens

<400> 348

```

Thr Gly Asp Ala Lys Gly Ala Val Thr Arg Gly Phe Ile Gly Ser Gly
1      5      10      15
Lys Val Val Thr Ala Ala Ala Val Ile Met Ile Ser Val Phe Val Phe
      20      25      30
Phe Ile Pro Glu Gly Met Asn Ala Ile Lys Glu Ile Ala Leu Ala Leu
      35      40      45
Ala Val Gly Ile Leu Thr Asp Ala Phe Leu Val Arg Met Thr Leu Val

```

```

      50              55              60
Pro Ala Val Met Ala Leu Leu Gly Asp Lys Ala Trp Trp Leu Pro Gly
65              70              75              80
Trp Leu Asp Arg Arg Leu Pro Arg Leu Asp Ile Glu Gly Glu Gly Ile
      85              90              95
Thr His Glu Glu Lys Leu Ala Ala Trp Pro Thr Ala Asp His Thr Glu
      100              105              110
Ala Leu His Ala Glu Gly Ile Gly Val Glu Gly Leu Phe Glu Gly Leu
      115              120              125
Asp Leu His Val Glu Pro Arg Gln Val Gln Ala Val Val Gly Ser Gln
      130              135              140
Asn Ser Val Ser Ala Val Leu Leu Ala Ile Gly Gly Arg Leu Pro Leu
145              150              155              160
Asp His Gly Arg Met Arg Ser Gly Gly Leu Leu Leu Pro Glu Arg Ala
      165              170              175
Ser Arg Val Arg Arg Val Thr Trp Phe Leu Asp Ala
      180              185

```

<210> 349
 <211> 339
 <212> DNA
 <213> Homo sapiens

```

<400> 349
ntgctggcca cggataatga ccgtactctg cgtgatgtcg ttgccgctga ccctacccat
60
gagctcgggtt cggctaccgc tcatacgttt gcggacaatt tgccgttcct tcttaaactg
120
ctcgcggcag aagagccact atcgttgacg gctcatccca gtttggcgca agcacaggaa
180
gggtacgggc gggagaatcg caaaggggtg ccattagatg cccagaccg gaattaccac
240
gatcccaacc ataaaccgga gcttattgtt gggctgacgc gattccacgc actagccggc
300
ttccgtgaac cacaacgcac acttgagctt tttgacgcg
339

```

<210> 350
 <211> 113
 <212> PRT
 <213> Homo sapiens

```

<400> 350
Xaa Leu Ala Thr Asp Asn Asp Arg Thr Leu Arg Asp Val Val Ala Ala
1      5      10      15
Asp Pro Thr His Glu Leu Gly Ser Ala Thr Ala His Thr Phe Ala Asp
      20      25      30
Asn Leu Pro Phe Leu Leu Lys Leu Leu Ala Ala Glu Glu Pro Leu Ser
      35      40      45
Leu Gln Ala His Pro Ser Leu Ala Gln Ala Gln Glu Gly Tyr Gly Arg
      50      55      60
Glu Asn Arg Lys Gly Val Pro Leu Asp Ala Pro Asp Arg Asn Tyr His
65      70      75      80
Asp Pro Asn His Lys Pro Glu Leu Ile Val Gly Leu Thr Arg Phe His

```

85 90 95
 Ala Leu Ala Gly Phe Arg Glu Pro Gln Arg Thr Leu Glu Leu Phe Asp
 100 105 110
 Ala

<210> 351
 <211> 354
 <212> DNA
 <213> Homo sapiens

<400> 351
 gcgcgccccca gtgccgagac ccggggcttc aggagccggc cccgggagag aagagtgcgg
 60
 cggcgggacgg agaaaacaac tccaaagtgt gcgaaaggca ccgcccctac tcccgggctg
 120
 ccgcgcctc cccgccccca gccctggcat ccagagtacg ggtcgagccc gnggccatgg
 180
 agccccctgt gggaggcggc accagggagc ctgggccccg gggctccgcc gcgaccccat
 240
 cgggtagacc acagaagctc cgggaccctt ccggcacctc tggacagccc aggatgctgt
 300
 tggccaccn ntctcctcc tctccttgg aggcgtctg gcccatccag accg
 354

<210> 352
 <211> 118
 <212> PRT
 <213> Homo sapiens

<400> 352
 Ala Arg Pro Ser Ala Glu Thr Arg Gly Phe Arg Ser Arg Pro Arg Glu
 1 5 10 15
 Arg Arg Val Arg Arg Arg Thr Glu Lys Thr Thr Pro Lys Leu Ala Lys
 20 25 30
 Gly Thr Ala Pro Thr Pro Gly Leu Pro Pro Pro Pro Arg Pro Gln Pro
 35 40 45
 Trp His Pro Glu Tyr Gly Ser Ser Pro Xaa Pro Trp Ser Pro Pro Gly
 50 55 60
 Glu Ala Ala Pro Gly Ser Leu Gly Pro Gly Ala Pro Pro Arg Pro His
 65 70 75 80
 Arg Val Asp His Arg Ser Ser Gly Thr Leu Pro Ala Pro Leu Asp Ser
 85 90 95
 Pro Gly Cys Cys Trp Pro Pro Xaa Pro Pro Pro Pro Trp Arg Arg
 100 105 110
 Ser Gly Pro Ser Arg Pro
 115

<210> 353
 <211> 1469
 <212> DNA
 <213> Homo sapiens

<400> 353

nntcatgaag gcttgaactt gcgtgatctt cagcctgcgg acctggcggg tgacggcggt
60
attgagccgg tggacctcgt ggteggagat gtctctttta tctccttgac gatgatcctt
120
gaacccattt cagctgttgt cagccacac ggctcatgc tgttgctggg gaagcctcaa
180
tttgagggtg gttgcaaggc tttgggagcc catggcggtg tcacggacct ggccctgcgc
240
ttgcaggcca tcgcggtgt catggcagca gcggtagatt tgggttggcg tatgctgac
300
gagtgcgata gcccggtgcc cgggcaggat ggaaacgttg agcacttcgt cttgctggaa
360
cgtacgggtc ggtgacagac gtccgggcat atcatgggccc gctactgtgg tcttgtgaac
420
gacacgagcc cttcgagata cgttgctgct gtcacccatg ccacgcggga cgacgctttt
480
gacgcggctg ccgaattcat ctctgaaatg gcggggcgag acattgggtg cgcggttccg
540
gatgatcagg tgaagccgat gtcaagcaag ctgccaggga tcgatcttga aagcttggga
600
gagttcgccc acgaggcgga ggtggtcgtc gtctttggcg gcgacggcac gatcttgcga
660
gctgctgaat ggtcattacc tcgccacgtt cccatgattg gcgtcaacct tggccatgtc
720
ggttttctgg ctgagctgga gcgctccgat atggcggtac tagtgaacaa ggtgtgttcg
780
cgcgactaca ccgttgagga tcgcctcgtg cttaaaacca ccgtcaccga gcattccgga
840
caacaccgtt ggagttcttt tgccgtcaac gagttgtctc tggaaaaggc agccggcg
900
cgcgtgctcg acgttctggc gtctgtcgac gagttgccgg tgcaacgctg gagttgcgac
960
gggatcctgg tctcgacccc gaccggatcg acggcctacg cgttctcagc tggcgccccg
1020
gtcatgtggc ccgatctcga cgccatgctc atgggtgccg tgagcgctca cgctctcttt
1080
gtctgaccgc tggatcatgag ccagctgct cgagtggacc ttgacatcca gccagacggt
1140
tcagaatcgg cggttctgtg gtgcgacggg cgccgatcgt gcaccgtacg accgggggaa
1200
agaatcacgg tcgtccgcca tccgaccgt ctgcgcattg ctggtctggc cgcgagcccc
1260
ttcacatcgc gtctgggtcaa gaagtttgag ctcccgtca gcgggtggcg tcagggtcgt
1320
gaccgtcatc acctagagga gacttcgtga tacgtagtgt gcgaattcgt ggactcggcg
1380
tcatcgatga gacggtctc gaacctcat ccgcgctgac ggcagtcacc ggcgagaccg
1440
gcgcccggaaa gaccatggtg gtcaccggt
1469

<210> 354

<211> 318

<212> PRT

<213> Homo sapiens

<400> 354

```

Met Gly Arg Tyr Cys Gly Leu Val Asn Asp Thr Ser Pro Ser Arg Tyr
 1           5           10           15
Val Val Val Val Thr His Ala Thr Arg Asp Asp Ala Phe Asp Ala Ala
      20           25           30
Ala Glu Phe Ile Ser Glu Met Ala Gly Arg Asp Ile Gly Cys Ala Val
      35           40           45
Pro Asp Asp Gln Val Lys Pro Met Ser Ser Lys Leu Pro Gly Ile Asp
      50           55           60
Leu Glu Ser Leu Gly Glu Phe Ala His Glu Ala Glu Val Val Val Val
65           70           75           80
Phe Gly Gly Asp Gly Thr Ile Leu Arg Ala Ala Glu Trp Ser Leu Pro
      85           90           95
Arg His Val Pro Met Ile Gly Val Asn Leu Gly His Val Gly Phe Leu
      100          105          110
Ala Glu Leu Glu Arg Ser Asp Met Ala Asp Leu Val Asn Lys Val Cys
      115          120          125
Ser Arg Asp Tyr Thr Val Glu Asp Arg Leu Val Leu Lys Thr Thr Val
      130          135          140
Thr Glu His Ser Gly Gln His Arg Trp Ser Ser Phe Ala Val Asn Glu
145          150          155          160
Leu Ser Leu Glu Lys Ala Ala Arg Arg Arg Met Leu Asp Val Leu Ala
      165          170          175
Ser Val Asp Glu Leu Pro Val Gln Arg Trp Ser Cys Asp Gly Ile Leu
      180          185          190
Val Ser Thr Pro Thr Gly Ser Thr Ala Tyr Ala Phe Ser Ala Gly Gly
      195          200          205
Pro Val Met Trp Pro Asp Leu Asp Ala Met Leu Met Val Pro Leu Ser
      210          215          220
Ala His Ala Leu Phe Ala Arg Pro Leu Val Met Ser Pro Ala Ala Arg
225          230          235          240
Val Asp Leu Asp Ile Gln Pro Asp Gly Ser Glu Ser Ala Val Leu Trp
      245          250          255
Cys Asp Gly Arg Arg Ser Cys Thr Val Arg Pro Gly Glu Arg Ile Thr
      260          265          270
Val Val Arg His Pro Asp Arg Leu Arg Ile Ala Arg Leu Ala Ala Gln
      275          280          285
Pro Phe Thr Ser Arg Leu Val Lys Lys Phe Glu Leu Pro Val Ser Gly
      290          295          300
Trp Arg Gln Gly Arg Asp Arg His His Leu Glu Glu Thr Ser
305          310          315

```

<210> 355

<211> 558

<212> DNA

<213> Homo sapiens

<400> 355

```

nggatccac ctctggaat ggaaaccac ataccagttc tcttctcga tttgaatgcg
60
gatgacctca gtgccaatga gcagcttgtt ggcccccatg catccggcgt gaactccatc
120

```

ctgccaagg agcatggcag ccagtttttc tacctgcca tcataaagca cagtgatgat
 180
 gaggtttcag ccacagcctc ttgggattcc tcggtgcatg attctgttca cttgaatggg
 240
 gtcacaccac agaagtaaag gatttaccta attgtgaaaa ccacagttca actcagccac
 300
 cctgtgcta tggagttagt attacgaaaa cgaattgcag ccaatattta caacaacag
 360
 agtttcacgc agagtttgaa gaggagaata tccctgaaaa atatatttta ttctgtggt
 420
 gtaacctatg aaatagtatc caatatacca aaggcaactg aggagataga ggaccgggaa
 480
 acgctggctc tcctggcagc aaggagtga aacgaaggca catcagatgg gaagacgtac
 540
 attgagaagt acactcga
 558

<210> 356

<211> 186

<212> PRT

<213> Homo sapiens

<400> 356

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Xaa | Ile | Pro | Pro | Pro | Gly | Met | Glu | Thr | His | Ile | Pro | Val | Leu | Phe | Leu |
| 1 | | | | 5 | | | | | 10 | | | | 15 | | |
| Asp | Leu | Asn | Ala | Asp | Asp | Leu | Ser | Ala | Asn | Glu | Gln | Leu | Val | Gly | Pro |
| | | | 20 | | | | | 25 | | | | 30 | | | |
| His | Ala | Ser | Gly | Val | Asn | Ser | Ile | Leu | Pro | Lys | Glu | His | Gly | Ser | Gln |
| | | 35 | | | | | 40 | | | | | 45 | | | |
| Phe | Phe | Tyr | Leu | Pro | Ile | Ile | Lys | His | Ser | Asp | Asp | Glu | Val | Ser | Ala |
| | 50 | | | | | 55 | | | | 60 | | | | | |
| Thr | Ala | Ser | Trp | Asp | Ser | Ser | Val | His | Asp | Ser | Val | His | Leu | Asn | Gly |
| 65 | | | | | 70 | | | | 75 | | | | 80 | | |
| Val | Thr | Pro | Gln | Asn | Glu | Arg | Ile | Tyr | Leu | Ile | Val | Lys | Thr | Thr | Val |
| | | | 85 | | | | | | 90 | | | | 95 | | |
| Gln | Leu | Ser | His | Pro | Ala | Ala | Met | Glu | Leu | Val | Leu | Arg | Lys | Arg | Ile |
| | | | 100 | | | | | 105 | | | | 110 | | | |
| Ala | Ala | Asn | Ile | Tyr | Asn | Lys | Gln | Ser | Phe | Thr | Gln | Ser | Leu | Lys | Arg |
| | | 115 | | | | 120 | | | | | 125 | | | | |
| Arg | Ile | Ser | Leu | Lys | Asn | Ile | Phe | Tyr | Ser | Cys | Gly | Val | Thr | Tyr | Glu |
| | 130 | | | | 135 | | | | | 140 | | | | | |
| Ile | Val | Ser | Asn | Ile | Pro | Lys | Ala | Thr | Glu | Glu | Ile | Glu | Asp | Arg | Glu |
| 145 | | | | 150 | | | | | 155 | | | | 160 | | |
| Thr | Leu | Ala | Leu | Leu | Ala | Ala | Arg | Ser | Glu | Asn | Glu | Gly | Thr | Ser | Asp |
| | | | 165 | | | | | 170 | | | | | 175 | | |
| Gly | Lys | Thr | Tyr | Ile | Glu | Lys | Tyr | Thr | Arg | | | | | | |
| | | 180 | | | | | | 185 | | | | | | | |

<210> 357

<211> 323

<212> DNA

<213> Homo sapiens

<400> 357

acgcgtgcgt gtgttggtgt agtcgggtgt gtgcatgcgt gtgggtgtgc agcaggtggg
 60
 gtacgatcag gctgaaggct gatcaggcac aaggctctgg gggagagccc tggttccagc
 120
 cctggggtca gagcagcagg ggccagaaag acggcagggg tgagcactgc acccgctggg
 180
 cagggcaggg ccacagaagg cagggcatgg aggccacgtg aagggttga cagagtggat
 240
 ggatgtctcc ggaagcacct gcgtggccca gtcagcagga tcagactcgc atgtgtcagg
 300
 gtcacatgg gtcagcgagg atn
 323

<210> 358
 <211> 102
 <212> PRT
 <213> Homo sapiens

<400> 358
 Met Val Thr Leu Thr His Ala Ser Leu Ile Leu Leu Thr Gly Pro Arg
 1 5 10 15
 Arg Cys Phe Arg Arg His Pro Ser Thr Leu Ser Ser Pro Ser Arg Gly
 20 25 30
 Leu His Ala Leu Pro Ser Val Ala Leu Pro Cys Pro Ala Gly Ala Val
 35 40 45
 Leu Thr Pro Ala Val Phe Leu Ala Pro Ala Ala Leu Thr Pro Gly Leu
 50 55 60
 Glu Pro Gly Leu Ser Pro Arg Ala Leu Cys Leu Ile Ser Leu Gln Pro
 65 70 75 80
 Asp Arg Thr Pro Pro Ala Ala His Pro His Ala Cys Thr His Pro Thr
 85 90 95
 His Thr Thr His Ala Arg
 100

<210> 359
 <211> 265
 <212> DNA
 <213> Homo sapiens

<400> 359
 acgcgtaccg acaagcgcgc ggtgatggcc gaccttcgcg aatcgggcgc aatcgagcag
 60
 gatgcggaca tgatcgtctt catctaccgc gacgattact acaacaagga aaattcgccg
 120
 gacaaggggc tggccgagat catcatcggc aagcatcggg ggggccccac cggctcgtgc
 180
 aagctgaagt tcttcggcga gtacaccggt ttcgacaacc tggcccacaa ctcggttggt
 240
 tcgttcgaat aacggatgat tccgg
 265

<210> 360
 <211> 83
 <212> PRT

<213> Homo sapiens

<400> 360

```

Thr Arg Thr Asp Lys Arg Pro Val Met Ala Asp Leu Arg Glu Ser Gly
 1             5             10             15
Ala Ile Glu Gln Asp Ala Asp Met Ile Val Phe Ile Tyr Arg Asp Asp
      20             25             30
Tyr Tyr Asn Lys Glu Asn Ser Pro Asp Lys Gly Leu Ala Glu Ile Ile
      35             40             45
Ile Gly Lys His Arg Gly Gly Pro Thr Gly Ser Cys Lys Leu Lys Phe
      50             55             60
Phe Gly Glu Tyr Thr Arg Phe Asp Asn Leu Ala His Asn Ser Val Gly
65             70             75             80
Ser Phe Glu

```

<210> 361

<211> 453

<212> DNA

<213> Homo sapiens

<400> 361

```

gctttgcagg aggaaatctc tatctctggc tgcaagatga ggctgagcta cctgagcagc
60
cggacccctg gctacaaatc tgtcctgagg atcagcctca cccacccgac catccccttc
120
aacctcatga aggtgcacct catggtagcg gtggagggcc gcctcttcag gaagtggttc
180
gctgcagccc cagacctgtc ctattatttc atttgggaca agacagacgt ctacaaccag
240
aaggtgtttg ggctttcaga agcctttggt tccgtggggt atgaatatga atcctgccca
300
gatctaattc tgtgggaaaa aagaacaaca gtgctgcagg gctatgaaat tgacgcgtcc
360
aagcttggag gatggagcct agacaaacat catgcctca acattcaaag tggcatcctg
420
caciaaggga atgngagaa ccagtttgtg tct
453

```

<210> 362

<211> 151

<212> PRT

<213> Homo sapiens

<400> 362

```

Ala Leu Gln Glu Glu Ile Ser Ile Ser Gly Cys Lys Met Arg Leu Ser
 1             5             10             15
Tyr Leu Ser Ser Arg Thr Pro Gly Tyr Lys Ser Val Leu Arg Ile Ser
      20             25             30
Leu Thr His Pro Thr Ile Pro Phe Asn Leu Met Lys Val His Leu Met
      35             40             45
Val Ala Val Glu Gly Arg Leu Phe Arg Lys Trp Phe Ala Ala Ala Pro
      50             55             60
Asp Leu Ser Tyr Tyr Phe Ile Trp Asp Lys Thr Asp Val Tyr Asn Gln

```

```

65          70          75          80
Lys Val Phe Gly Leu Ser Glu Ala Phe Val Ser Val Gly Tyr Glu Tyr
          85          90          95
Glu Ser Cys Pro Asp Leu Ile Leu Trp Glu Lys Arg Thr Thr Val Leu
          100          105          110
Gln Gly Tyr Glu Ile Asp Ala Ser Lys Leu Gly Gly Trp Ser Leu Asp
          115          120          125
Lys His His Ala Leu Asn Ile Gln Ser Gly Ile Leu His Lys Gly Asn
          130          135          140
Gly Glu Asn Gln Phe Val Ser
145          150

```

<210> 363

<211> 502

<212> DNA

<213> Homo sapiens

<400> 363

```

gggtacaaaa aagtttgcca cagtattcac actccaggtc tccataaacc ttccagatcc
60
gctcacacaa gctggtgttc atttgcttct tctgtaaact gttcaggacc ttcataaaaag
120
cggtgatgcc tgaccggtgc tcaggggcag ctttgcaaga gtcaggctga tgtgtgatgg
180
tgtccccacc accagctact ggagggagga ggtctgaggc ctcagctggg tttgacctga
240
gacacctgct gggatctggg tcaccagctg aaagcacagc catgttctgc ctttccctta
300
gggggctctg ggcgccatgg ctttcctgat ctgaccagc actctgggccc ttggacagca
360
gtagtgtgat cacttcacct tgcgtctgga ctgagcttct gtgctgcatg tctgggggct
420
tctcaggagc agcatgagcc tctgaggagg aggtatcatt tttcaacaaa aaatcatctg
480
aaaccacctc ttgagaatgc ag
502

```

<210> 364

<211> 136

<212> PRT

<213> Homo sapiens

<400> 364

```

Met Gln His Arg Ser Ser Val Gln Thr Gln Gly Glu Val Ile Thr Leu
1          5          10          15
Leu Leu Ser Lys Ala Gln Ser Ala Gly Ser Asp Gln Glu Ser His Gly
          20          25          30
Ala Gln Ser Pro Leu Gly Glu Gly Gln Asn Met Ala Val Leu Ser Ala
          35          40          45
Gly Asp Pro Asp Pro Ser Arg Cys Leu Arg Ser Asn Pro Ala Glu Ala
          50          55          60
Ser Asp Leu Leu Pro Pro Val Ala Gly Gly Gly Asp Thr Ile Thr His
65          70          75          80
Gln Pro Asp Ser Cys Lys Ala Ala Pro Glu His Arg Ser Gly Ile Thr

```

| | | | | | |
|---|-----|--|-----|--|-----|
| | 85 | | 90 | | 95 |
| Ala Phe Met Lys Val Leu Asn Ser Leu Gln Lys Lys Gln Met Asn Thr | | | | | |
| | 100 | | 105 | | 110 |
| Ser Leu Cys Glu Arg Ile Trp Lys Val Tyr Gly Asp Leu Glu Cys Glu | | | | | |
| | 115 | | 120 | | 125 |
| Tyr Cys Gly Lys Leu Phe Trp Tyr | | | | | |
| | 130 | | 135 | | |

<210> 365

<211> 333

<212> DNA

<213> Homo sapiens

<400> 365

atctcaacgg atgcatccat caaggagatg atccccccag gtgctcttgt tatgctcaca
60
ccactgatcg ttgggattct atttgggggt gagaccctct ctggagtcct tgctgggtgcc
120
cttgctctctg gtgttcagat tgccatttct gcatccaaca ctgggtgggtgc ctgggacaac
180
gccaagaagt acattgaggc tggagtttca gagcatgccca ggacccttgg cccaaaaggt
240
tctgaccctc acaaggcggc tgtcattggt gacaccattg gagatcctct caaggacacg
300
tctggccctt ccctcaacat cctcatcaag ctt
333

<210> 366

<211> 111

<212> PRT

<213> Homo sapiens

<400> 366

| | | | | | |
|---|-----|----|-----|----|-----|
| Ile Ser Thr Asp Ala Ser Ile Lys Glu Met Ile Pro Pro Gly Ala Leu | | | | | |
| 1 | 5 | | 10 | | 15 |
| Val Met Leu Thr Pro Leu Ile Val Gly Ile Leu Phe Gly Val Glu Thr | | | | | |
| | 20 | | 25 | | 30 |
| Leu Ser Gly Val Leu Ala Gly Ala Leu Val Ser Gly Val Gln Ile Ala | | | | | |
| | 35 | | 40 | | 45 |
| Ile Ser Ala Ser Asn Thr Gly Gly Ala Trp Asp Asn Ala Lys Lys Tyr | | | | | |
| | 50 | | 55 | | 60 |
| Ile Glu Ala Gly Val Ser Glu His Ala Arg Thr Leu Gly Pro Lys Gly | | | | | |
| 65 | | 70 | | 75 | 80 |
| Ser Asp Pro His Lys Ala Ala Val Ile Gly Asp Thr Ile Gly Asp Pro | | | | | |
| | 85 | | 90 | | 95 |
| Leu Lys Asp Thr Ser Gly Pro Ser Leu Asn Ile Leu Ile Lys Leu | | | | | |
| | 100 | | 105 | | 110 |

<210> 367

<211> 381

<212> DNA

<213> Homo sapiens

<400> 367

gcgttcgtcg cactacccgg cggcggcgga acccttgacg agctactcga agcatggaca
 60
 tggcagcagc tcggtgtaca cagcaaaccg gtgngccttg tacgactcga cnncttctgg
 120
 gcaccgctga ccgcgctact caaccacatg accatcgaaa gcttcattcg ccctgaggac
 180
 cgcgcctcgc tcgtgatcgc cgataccata catcagctga tggccgatct tgagggatgg
 240
 accccaccac caccgaagtg gcgctcgtga catagaacaa atgattctga ctatggctca
 300
 ttgacatctg cgcagcggt actagctcca ttgacttcaa atcgggcctt ggccgaggct
 360
 cngttcaggt ggcccgaat g
 381

<210> 368

<211> 89

<212> PRT

<213> Homo sapiens

<400> 368

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Ala | Phe | Val | Ala | Leu | Pro | Gly | Gly | Gly | Gly | Thr | Leu | Asp | Glu | Leu | Leu |
| 1 | | | | 5 | | | | 10 | | | | | | 15 | |
| Glu | Ala | Trp | Thr | Trp | Gln | Gln | Leu | Gly | Val | His | Ser | Lys | Pro | Val | Xaa |
| | | | 20 | | | | | 25 | | | | | 30 | | |
| Leu | Val | Arg | Leu | Asp | Xaa | Phe | Trp | Ala | Pro | Leu | Thr | Ala | Leu | Leu | Asn |
| | | 35 | | | | | 40 | | | | | 45 | | | |
| His | Met | Thr | Ile | Glu | Ser | Phe | Ile | Arg | Pro | Glu | Asp | Arg | Ala | Ser | Leu |
| | 50 | | | | | 55 | | | | 60 | | | | | |
| Val | Ile | Ala | Asp | Thr | Ile | His | Gln | Leu | Met | Ala | Asp | Leu | Glu | Gly | Trp |
| 65 | | | | 70 | | | | | 75 | | | | | 80 | |
| Thr | Pro | Pro | Pro | Pro | Lys | Trp | Arg | Ser | | | | | | | |
| | | | | 85 | | | | | | | | | | | |

<210> 369

<211> 313

<212> DNA

<213> Homo sapiens

<400> 369

gatacatgat cctctcatat cgcacacaca ccgctcccct ctgccgcaat tcgcagacaa
 60
 acttgcgagc gcttcacagc aagccgtcaa ggctgcttcc tgtgggctac cgatagtctc
 120
 gtacgcgagt tctcggacat caacgccaac gtcgggcaag atactgtcaa cgccatctac
 180
 acattctacg agcagcaagc gaccagtttc cttcgccagc tgaacgacct cccacccgaa
 240
 gagcttcccg acgtcatcga ggactttctc cgctgtcca ctgatgtcct tctttaccat
 300
 ttccagcaag ctt
 313

<210> 370

<211> 101
 <212> PRT
 <213> Homo sapiens

<400> 370
 Ser Ser His Thr Ala His Thr Pro Leu Pro Ser Ala Ala Ile Arg Arg
 1 5 10 15
 Gln Thr Cys Ala Gly Phe Thr Ala Ser Arg Gln Gly Cys Phe Leu Trp
 20 25 30
 Ala Thr Asp Ser Leu Val Arg Glu Phe Ser Asp Ile Asn Ala Asn Val
 35 40 45
 Gly Gln Asp Thr Val Asn Ala Ile Tyr Thr Phe Tyr Glu Gln Gln Ala
 50 55 60
 Thr Ser Phe Leu Arg Gln Leu Asn Asp Leu Pro Pro Glu Glu Leu Pro
 65 70 75 80
 Asp Val Ile Glu Asp Phe Phe Arg Leu Ser Thr Asp Val Leu Leu Tyr
 85 90 95
 His Phe Gln Gln Ala
 100

<210> 371
 <211> 380
 <212> DNA
 <213> Homo sapiens

<400> 371
 atgacgggtc acgtcatcct ggcgattcca caggtggtga cgatcatggat cggcctcatc
 60
 tgcacgccta ttggcacggg ctttatcaag ccgaacctct ccacgggtgt aggaggtctt
 120
 tacgatgacg gtgacccccg ccgcgatcag gggttctctgt acttctacat gtcgatcagt
 180
 attggtatctc tcttcgcgcc gatcgtcacc ggcctcctca aggaccatta cggctaccac
 240
 gtaggtttca ttgccgctgc taccggtatg gctctgggtc tgatcgctt cttccacggt
 300
 cgttccaaac tgcgtgagct cgccttcgac atcccgaatc cgttggtccc cggcgagggt
 360
 cgccggatgg tgctccgcgg
 380

<210> 372
 <211> 126
 <212> PRT
 <213> Homo sapiens

<400> 372
 Met Thr Gly His Val Ile Leu Ala Ile Pro Gln Val Val Thr Ser Trp
 1 5 10 15
 Ile Gly Leu Ile Cys Ile Ala Ile Gly Thr Gly Phe Ile Lys Pro Asn
 20 25 30
 Leu Ser Thr Val Val Gly Gly Leu Tyr Asp Asp Gly Asp Pro Arg Arg
 35 40 45
 Asp Gln Gly Phe Leu Tyr Phe Tyr Met Ser Ile Ser Ile Gly Ser Leu

```

      50              55              60
Phe Ala Pro Ile Val Thr Gly Leu Leu Lys Asp His Tyr Gly Tyr His
65              70              75              80
Val Gly Phe Ile Ala Ala Ile Gly Met Ala Leu Gly Leu Ile Ala
      85              90              95
Phe Phe His Gly Arg Ser Lys Leu Arg Glu Leu Ala Phe Asp Ile Pro
      100              105              110
Asn Pro Leu Ala Pro Gly Glu Gly Arg Arg Met Val Leu Arg
      115              120              125

```

<210> 373
 <211> 475
 <212> DNA
 <213> Homo sapiens

```

<400> 373
acatgttgga aaaattgcct cccactctgg tgctacaggt atgaatctca gccacagtga
60
tgactgtggc agctacaggc ctgatgaaca cccaccaag aaaaggagca tcatgtgcct
120
gcttctctct ggttcctaaa tcctttggcc aaacattttc cccacaaccc tccactccag
180
ttggctggtc actgcctctc agaaagaagt cccagggtccc tgtcagcccc agagcgctcg
240
catggactct gccactgtc cctttccaac acggaggccc ccaattctgg ggaccctac
300
accctaccct gtaccaccac atccccatgc ctgctccaga cagcactaac ctcccatgac
360
agtgggacca aagcagttct taaaggcca atccactcag ttcttaaatg aaaaacagtt
420
gcccatgagt ccccccaaa gacgtccgca catatgccaa acattcgggtg tgcac
475

```

<210> 374
 <211> 109
 <212> PRT
 <213> Homo sapiens

```

<400> 374
Met Gly Met Trp Trp Tyr Arg Val Gly Cys Arg Gly Pro Gln Asn Trp
1      5      10      15
Gly Pro Pro Cys Trp Lys Gly Thr Val Gly Arg Val His Ala Gly Ala
      20      25      30
Leu Gly Leu Thr Gly Thr Trp Asp Phe Phe Leu Arg Gly Ser Asp Gln
      35      40      45
Pro Thr Gly Val Glu Gly Cys Gly Glu Asn Val Trp Pro Lys Asp Leu
      50      55      60
Gly Thr Arg Glu Lys Gln Ala His Asp Ala Pro Phe Leu Gly Gly Val
65      70      75      80
Phe Ile Arg Pro Val Ala Ala Thr Val Ile Thr Val Ala Glu Ile His
      85      90      95
Thr Cys Ser Thr Arg Val Gly Gly Asn Phe Ser Asn Met
      100      105

```

<210> 375
 <211> 332
 <212> DNA
 <213> Homo sapiens

<400> 375
 nnacgcgtcg cctccacctc gaaacccgcc ggcggtcggt ttttcaccat ggccgaccgc
 60
 aaggcccaag ttgcgacggt cacggacacg ctgtatttca cgccgtcgca atgggatgga
 120
 tgcattggcac ggatgcgtgg ggataagata tcagcactga agtggaatca gatgcagatg
 180
 gcggcatgct ccttcatagc ggcagtgggt gcgaagctgg gctgcccgc ggcactatg
 240
 ggcacggcgc agctgctgta ccagcgtttc catctatttc atgcgccgac tgagttttcg
 300
 ttacatgagg tggctttgac gtgtctcttc ac
 332

<210> 376
 <211> 110
 <212> PRT
 <213> Homo sapiens

<400> 376
 Xaa Arg Val Ala Ser Thr Ser Lys Pro Ala Gly Gly Arg Phe Phe Thr
 1 5 10 15
 Met Ala Asp Arg Lys Ala Gln Val Ala Thr Val Thr Asp Thr Leu Tyr
 20 25 30
 Phe Thr Pro Ser Gln Trp Asp Gly Cys Met Ala Arg Met Arg Gly Asp
 35 40 45
 Lys Ile Ser Ala Leu Lys Trp Asn Gln Met Gln Met Ala Ala Cys Ser
 50 55 60
 Phe Ile Ala Ala Val Gly Ala Lys Leu Gly Cys Pro Gln Arg Thr Met
 65 70 75 80
 Gly Thr Ala Gln Leu Tyr Gln Arg Phe His Leu Phe His Ala Pro
 85 90 95
 Thr Glu Phe Ser Leu His Glu Val Ala Leu Thr Cys Leu Phe
 100 105 110

<210> 377
 <211> 369
 <212> DNA
 <213> Homo sapiens

<400> 377
 cgcggtgccag gtatgtcaac tgatctgtcg gatatttccg aggttgagta ccgtcaactg
 60
 aggctggaac gagggtggtg gtgttcgggtg tggactcagg gaactgccgc agacgccgag
 120
 aacgctatgg cggagctgaa agcccttgct gaaacggcgg gatctcaggt actcgaagct
 180
 gtcattgcaac gtcggactac cccggatccg gcgacgtaca ttggttcggg caaggtgggt
 240

gagcttgccg aggtggtgcg ggcgactggt gccgatactg tcatttgtga cggatgaactt
 300
 gacgccgctc agttgcgcaa cctcgaggat cgggtcaagn gcaaagttgt ggaccggtcg
 360
 gtctgattc
 369

<210> 378
 <211> 121
 <212> PRT
 <213> Homo sapiens

<400> 378
 Arg Val Pro Gly Met Ser Thr Asp Leu Ser Asp Ile Ser Glu Val Glu
 1 5 10 15
 Tyr Arg Gln Leu Arg Leu Glu Arg Val Val Leu Cys Ser Val Trp Thr
 20 25 30
 Gln Gly Thr Ala Ala Asp Ala Glu Asn Ala Met Ala Glu Leu Lys Ala
 35 40 45
 Leu Ala Glu Thr Ala Gly Ser Gln Val Leu Glu Ala Val Met Gln Arg
 50 55 60
 Arg Thr Thr Pro Asp Pro Ala Thr Tyr Ile Gly Ser Gly Lys Val Ala
 65 70 75 80
 Glu Leu Ala Glu Val Val Arg Ala Thr Gly Ala Asp Thr Val Ile Cys
 85 90 95
 Asp Gly Glu Leu Asp Ala Ala Gln Leu Arg Asn Leu Glu Asp Arg Val
 100 105 110
 Lys Xaa Lys Val Val Asp Arg Ser Val
 115 120

<210> 379
 <211> 408
 <212> DNA
 <213> Homo sapiens

<400> 379
 acgcgttact taaacttata tgtaaataat aaattcatta tttctagttg gtttaggtact
 60
 atgggctgtg gtttaccagg tgctatggca gctaaaattg cttatccaaa ccgtcaagca
 120
 gtagctatca caggcgacgg tgcgttccaa atggtaatgc aagactttgc tacagctggt
 180
 caatataact taccaatgac aatctttgta ttaaataaca aacaattgtc attcattaaa
 240
 tatgaacaac aagctgctgg tgaattagag tatgccattg atttctctga tatggatcat
 300
 gctaaatttg ctgaagctgc tgggtggtaaa ggctatgttg tgagagatgt aagtcgtctt
 360
 gacgacatcg ttgaagaggc aatgggtcaa gatgttccaa caatcggt
 408

<210> 380
 <211> 136
 <212> PRT

<213> Homo sapiens

<400> 380

```

Thr Arg Tyr Leu Asn Leu Ser Val Asn Asn Lys Phe Ile Ile Ser Ser
 1           5           10           15
Trp Leu Gly Thr Met Gly Cys Gly Leu Pro Gly Ala Met Ala Ala Lys
      20           25           30
Ile Ala Tyr Pro Asn Arg Gln Ala Val Ala Ile Thr Gly Asp Gly Ala
      35           40           45
Phe Gln Met Val Met Gln Asp Phe Ala Thr Ala Val Gln Tyr Asn Leu
      50           55           60
Pro Met Thr Ile Phe Val Leu Asn Asn Lys Gln Leu Ser Phe Ile Lys
      65           70           75           80
Tyr Glu Gln Gln Ala Ala Gly Glu Leu Glu Tyr Ala Ile Asp Phe Ser
      85           90           95
Asp Met Asp His Ala Lys Phe Ala Glu Ala Ala Gly Gly Lys Gly Tyr
      100          105          110
Val Val Arg Asp Val Ser Arg Leu Asp Asp Ile Val Glu Glu Ala Met
      115          120          125
Ala Gln Asp Val Pro Thr Ile Val
      130          135

```

<210> 381

<211> 613

<212> DNA

<213> Homo sapiens

<400> 381

```

naccgctcat aggcgggccc agtgaagac caccgaaca cagttggtg agatccgct
60
tgagggaag gtctgcgcg tcccgcaaa tctggtcaag gcctaccact ctgggctgat
120
cgacgtcgag gactgaaccc tgggagcctg ggcggtccag catgactgct caggctcatt
180
accaaaacgc gtcgatcccg taggggtgtc gtcagtagca agcccgaagt gaccctgccc
240
gattccgccc ccgacgacct cgtcgttgag gacatcacca tcggcgacgg cctgaagcg
300
tccgctggca acctcgtcga agtgcactac gtcggcgtgg ccttaagcaa tggctcgtgag
360
ttcgattctt cctggaaccg cggggagccg ctgaccttcc aactaggggc tggccagggtg
420
atccccgagt gggatgaagg tgtccaaggt atgaaggctg gtggacgacg caaactcgtc
480
atccccacc accttgctta cgggccgcaa ggaatctccg gtgtgatcgc tggcggtgag
540
acgctggtct tcgtctgcga ccttgtaaac atcatctgac gtgacccccg ctcaagcagt
600
cttcgcgccc ggg
613

```

<210> 382

<211> 137

<212> PRT

<213> Homo sapiens

<400> 382

```

Leu Leu Arg Leu Ile Thr Lys Thr Arg Arg Ser Arg Arg Val Val Val
1           5           10           15
Met Ser Lys Pro Glu Val Thr Leu Pro Asp Ser Ala Pro Asp Asp Leu
20           25           30
Val Val Glu Asp Ile Thr Ile Gly Asp Gly Pro Glu Ala Ser Ala Gly
35           40           45
Asn Leu Val Glu Val His Tyr Val Gly Val Ala Leu Ser Asn Gly Arg
50           55           60
Glu Phe Asp Ser Ser Trp Asn Arg Gly Glu Pro Leu Thr Phe Gln Leu
65           70           75           80
Gly Ala Gly Gln Val Ile Pro Glu Trp Asp Glu Gly Val Gln Gly Met
85           90           95
Lys Val Gly Gly Arg Arg Lys Leu Val Ile Pro His His Leu Ala Tyr
100          105          110
Gly Pro Gln Gly Ile Ser Gly Val Ile Ala Gly Gly Glu Thr Leu Val
115          120          125
Phe Val Cys Asp Leu Val Asn Ile Ile
130          135

```

<210> 383

<211> 352

<212> DNA

<213> Homo sapiens

<400> 383

```

nggagcaaca cctggtcctt gggaatgaag tgtaggagtt gcatttgctg aggttggtgt
60
ttgccaaaga gatgccagct tcttcgaact actgctgtgc aactcttcat gttcaaaacc
120
cagttttctg tttttcacac ctgaacatac acccccctgc agttgggtgg ctccccggtt
180
accagctggg ctctatctac agagagagca atggcttccc ttcccttgaa ggaagtctca
240
ccctcacaag gacacttgat ccgctgcaaa gcagaaagtg tgcggaccct ttgggaaggg
300
cgttcttttc ttgttttagaa cctaggattc tgtttttccc aaacaggatc an
352

```

<210> 384

<211> 93

<212> PRT

<213> Homo sapiens

<400> 384

```

Met Pro Ala Ser Ser Asn Tyr Cys Cys Ala Thr Leu His Val Gln Asn
1           5           10           15
Pro Val Phe Cys Phe Ser His Leu Asn Ile His Pro Pro Ala Val Gly
20           25           30
Trp Leu Pro Arg Tyr Gln Leu Gly Ser Ile Tyr Arg Glu Ser Asn Gly
35           40           45
Phe Pro Ser Leu Glu Gly Ser Leu Thr Leu Thr Arg Thr Leu Asp Pro

```

50 55 60
 Leu Gln Ser Arg Lys Cys Ala Asp Pro Leu Gly Arg Ala Phe Phe Ser
 65 70 75 80
 Cys Leu Glu Pro Arg Ile Leu Phe Phe Pro Asn Arg Ile
 85 90

<210> 385
 <211> 342
 <212> DNA
 <213> Homo sapiens

<400> 385
 gccggcgcca cgaaatgcaa aatgcgccct tcaccggacg ccagggtgat cgagccgcca
 60
 gcacctcggg caatgtcctg ggcctgactg gcacacgcaa tcaaagcgag caacaacaca
 120
 caaaaacgca tcatgaggca gacgccaggg aagtacaga agccgcagca ggcgcgcggc
 180
 gattggaaat atcggtgagg ctaatggtca ccagcgcttg caggttgat tcggtggcca
 240
 attcgcgga cgacagcacc gccagttcca gctcgccgcg cagcaccagg cgacgcaagc
 300
 tgcggcgcaa ctccgggtgc accaacaaca ccgcactgtt ca
 342

<210> 386
 <211> 109
 <212> PRT
 <213> Homo sapiens

<400> 386
 Met Gln Asn Ala Pro Phe Thr Gly Arg Gln Val Asp Arg Ala Ala Ser
 1 5 10 15
 Thr Ser Gly Asn Val Leu Gly Leu Thr Gly Thr Arg Asn Gln Ser Glu
 20 25 30
 Gln Gln His Thr Lys Thr His His Glu Ala Asp Ala Arg Glu Val Thr
 35 40 45
 Glu Ala Ala Ala Gly Ala Arg Arg Leu Glu Ile Ser Val Arg Leu Met
 50 55 60
 Val Thr Ser Ala Cys Arg Leu Tyr Ser Val Ala Asn Ser Arg Asn Asp
 65 70 75 80
 Ser Thr Ala Ser Ser Ser Ser Pro Arg Ser Thr Arg Arg Arg Lys Leu
 85 90 95
 Arg Arg Asn Ser Gly Cys Thr Asn Asn Thr Ala Leu Phe
 100 105

<210> 387
 <211> 379
 <212> DNA
 <213> Homo sapiens

<400> 387
 acgcgtgacg cgccggcatc ggaagcgttg actgcagaga agaccgcgca cgtggctgtg
 60

ggacgtgctg gcacgtctga catggtgcgt ggacccgcct tctcttcgcc tgcgcatgcc
 120
 atgcaagagg agcttgacaa tgtgcgtgat ctgcgccatg cgcggcagca agcgctcgat
 180
 gctgttcggt ccgagctgct cgaagcgcag caagcatgtg cctcgtgccca gctgcagctg
 240
 cagcatgtgc cagatgatcg tgtgcgagcg catcccatat accaggcgct ccatgcggac
 300
 gttgcttaca tgcagcaaga acttgatcac gtacgagacg cattggcttc ggcagaatct
 360
 gagaatgcga gcctgcgcg
 379

<210> 388

<211> 114

<212> PRT

<213> Homo sapiens

<400> 388

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Arg | Leu | Val | Arg | Asp | Gln | Val | Leu | Ala | Ala | Cys | Lys | Gln | Arg | Pro |
| 1 | | | | 5 | | | | | 10 | | | | | 15 | |
| His | Gly | Ala | Pro | Gly | Ile | Trp | Asp | Ala | Leu | Ala | His | Asp | His | Leu | Ala |
| | | 20 | | | | | | 25 | | | | | 30 | | |
| His | Ala | Ala | Ala | Ala | Ala | Gly | Thr | Arg | His | Met | Leu | Ala | Ala | Leu | Arg |
| | | 35 | | | | | 40 | | | | | 45 | | | |
| Ala | Ala | Arg | Asn | Glu | Gln | His | Arg | Ala | Leu | Ala | Ala | Ala | His | Gly | Arg |
| | 50 | | | | | 55 | | | | | 60 | | | | |
| Asp | His | Ala | His | Cys | Gln | Ala | Pro | Leu | Ala | Trp | His | Ala | Gln | Ala | Lys |
| 65 | | | | | 70 | | | | | 75 | | | | 80 | |
| Arg | Arg | Arg | Val | His | Ala | Pro | Cys | Gln | Thr | Cys | Gln | His | Val | Pro | Gln |
| | | | 85 | | | | | 90 | | | | | 95 | | |
| Pro | Arg | Ala | Arg | Ser | Ser | Leu | Gln | Ser | Thr | Leu | Pro | Met | Pro | Ala | Arg |
| | | | 100 | | | | | 105 | | | | | 110 | | |
| His | Ala | | | | | | | | | | | | | | |

<210> 389

<211> 382

<212> DNA

<213> Homo sapiens

<400> 389

ngatggccga ctgtcccaact gtcagtagcg gaagctcgcc gtcgagtcgg tccacgtccg
 60
 ggcctcccaac gtgctccgca accctccgaa gcgatgacct ggccccgggg cggaacagag
 120
 gtattgcgtt tggagacgct tgggggtcaat tacggccagg tgcgcgccgt cgatgccttg
 180
 acgaccaccg tagagcgcg caccatcacc tgctcatgg gtcgaaatgg atcaggcaag
 240
 tcgtctctga tgtgggcat ccaaggggca acaaagtcct caggagggt actggtcaac
 300
 cagcagggtt cttgggctga cccccgaaa gccgacgccg cgaccgctcg acgaatggtg
 360

agcttagtcc cgcagtcagc cn
382

<210> 390
<211> 127
<212> PRT
<213> Homo sapiens

<400> 390
Xaa Trp Pro Thr Val Pro Leu Ser Val Arg Glu Ala Arg Arg Arg Val
1 5 10 15
Gly Pro Arg Pro Gly Leu Pro Arg Ala Pro Gln Pro Ser Glu Ala Met
20 25 30
Thr Trp Pro Gly Gly Gly Asn Glu Val Leu Arg Leu Glu Thr Leu Gly
35 40 45
Val Asn Tyr Gly Gln Val Arg Ala Val Asp Ala Leu Thr Thr Thr Val
50 55 60
Glu Arg Gly Thr Ile Thr Cys Leu Met Gly Arg Asn Gly Ser Gly Lys
65 70 75 80
Ser Ser Leu Met Trp Ala Ile Gln Gly Ala Thr Lys Ser Ser Gly Arg
85 90 95
Val Leu Val Asn His Glu Gly Ser Trp Ala Asp Pro Arg Lys Ala Asp
100 105 110
Ala Ala Thr Ala Arg Arg Met Val Ser Leu Val Pro Gln Ser Ala
115 120 125

<210> 391
<211> 456
<212> DNA
<213> Homo sapiens

<400> 391
nnacgcgttg ccgctctgtg aggcgcctat cacggtgaca ctctcggtgc tatgagcgtg
60
tgcgacccta tcggtggcat gcacgccttg ttcagcgact ctattcccca gcagatcttc
120
ctgcccgcgc cctccttctt tcgccgccga cgaggccgac gtggagacgt ggtgcagcga
180
ggccgatgaa tcctggacac ccaccgcgac gacctggccg ggatcattgt cgagcccatc
240
ttgcaaggag ccggaggcat gtggccgtgg tctccgtcct gtctgaagca cctgcgccgt
300
cgtgctgatg aacttgacct agttcttata gccgacgagg tcgctactgg atttgggcgg
360
actggcaaac ttttcgcatg cgagtgggccc gatatcggtc ctgacatcat ggtgggtggg
420
aaatccatga ctggcggata cctgacccag tcggcc
456

<210> 392
<211> 55
<212> PRT
<213> Homo sapiens

<400> 392

Gly Ala Tyr His Gly Asp Thr Leu Gly Ala Met Ser Val Cys Asp Pro
 1 5 10 15
 Ile Gly Gly Met His Ala Xaa Phe Ser Asp Ser Ile Pro Gln Gln Ile
 20 25 30
 Phe Leu Pro Ala Pro Ser Phe Phe Arg Arg Arg Arg Gly Arg Arg Gly
 35 40 45
 Asp Val Val Gln Arg Gly Arg
 50 55

<210> 393

<211> 371

<212> DNA

<213> Homo sapiens

<400> 393

nacgcgttgc tcgtcattgg tggctactcg gcctacgaag gtatctacac catgatgact
 60
 gagcgggacc ggtacccggc tttccgtatt cgcacgggtg gcaccccggc ttctatcgac
 120
 aacaacctcc ccggttcgga actgtccatc ggcaccgaca ccgctctcaa cgtcatcgtc
 180
 gaggcgatgg acaagattaa ggagtcgggt atcgcgtcca gacgctgctt cgtcgtcgag
 240
 acgatgggtc gtgactgcgg atacctcgcg ttgatgtcgg gtatcgcagc tggcgtcgag
 300
 cggatctata ccaacgagga cggtatctcc ctggacgac tagccaacga cgtccattgg
 360
 ttgcgggagt c
 371

<210> 394

<211> 123

<212> PRT

<213> Homo sapiens

<400> 394

Xaa Ala Leu Leu Val Ile Gly Gly Tyr Ser Ala Tyr Glu Gly Ile Tyr
 1 5 10 15
 Thr Met Met Thr Glu Arg Asp Arg Tyr Pro Ala Phe Arg Ile Pro Thr
 20 25 30
 Val Cys Ile Pro Ala Ser Ile Asp Asn Asn Leu Pro Gly Ser Glu Leu
 35 40 45
 Ser Ile Gly Thr Asp Thr Ala Leu Asn Val Ile Val Glu Ala Met Asp
 50 55 60
 Lys Ile Lys Glu Ser Gly Ile Ala Ser Arg Arg Cys Phe Val Val Glu
 65 70 75 80
 Thr Met Gly Arg Asp Cys Gly Tyr Leu Ala Leu Met Ser Gly Ile Ala
 85 90 95
 Ala Gly Ala Glu Arg Ile Tyr Thr Asn Glu Asp Gly Ile Ser Leu Asp
 100 105 110
 Asp Leu Ala Asn Asp Val His Trp Leu Arg Glu
 115 120

<210> 395
 <211> 351
 <212> DNA
 <213> Homo sapiens

<400> 395
 gaattctagt tgggagattc attgaccaga cttttggaat aaacactagt catcatgcta
 60
 gcgacagggtg gtcttgtgca tggtagaaag gcagccaag cctatgtctc tgaaacctgc
 120
 tctcatttct gttttctact ttacgattta tggtatctca tactcccat gttgcctggt
 180
 ctccagtttt ttacttgtg ttatttccat tcttctatc ctgctcaatt tctgcctcag
 240
 ggcagaattg tgtccaacag ctcttaaatg cagcgagaa actgtgatgt taaaaacatc
 300
 ttgttatccg gcccacaaac atgttgcct tggtaactct tactggttg t
 351

<210> 396
 <211> 90
 <212> PRT
 <213> Homo sapiens

<400> 396
 Met Val Glu Arg Gln Ser Lys Pro Met Ser Leu Lys Pro Ala Leu Ile
 1 5 10 15
 Ser Val Phe Tyr Phe Thr Ile Tyr Val Ile Ser Tyr Ser Pro Cys Cys
 20 25 30
 Leu Phe Ser Ser Phe Phe Thr Cys Val Ile Ser Ile Leu Leu Phe Leu
 35 40 45
 Leu Asn Phe Cys Leu Arg Ala Glu Leu Cys Pro Thr Ala Leu Lys Cys
 50 55 60
 Ser Ala Glu Thr Val Met Leu Lys Thr Ser Cys Tyr Pro Ala Pro Lys
 65 70 75 80
 His Val Val Leu Gly Asn Ser Tyr Trp Phe
 85 90

<210> 397
 <211> 483
 <212> DNA
 <213> Homo sapiens

<400> 397
 gccgtcatta aagagatcac ccctctctc caacctgggtg atgtcctcgt cgacggtggt
 60
 aatgcttatt ttggtgatac ccgccgccgt gaggaggaaa tacgtccac cggcattcac
 120
 tatgttggtg ctggcatctc cggtggggga gtcggggccc tgagggtccc atcaattatg
 180
 cctggcgggg ttaaggaatc ttacgaaatc atcggaaccg tcttagaaaa aatctccgcc
 240
 cacgtcgacg gtgaacctg ctgcgcatgg atgggtactg acggcgccgg acacttcgtc
 300

aagatggtcc ataatggcat cgagtacgcc gatatgcagt tcattggcga ggcgcccttc
 360
 ctttttgcn tgcccgcgg tttgaccaat gctgaggccg ccgatgcctt cgagtcgtgg
 420
 aaccatggcg acctcaattc ctacctgctc gaaatcactt ctcgggtact gcgtgccaag
 480
 gat
 483

<210> 398
 <211> 161
 <212> PRT
 <213> Homo sapiens

<400> 398
 Ala Val Ile Lys Glu Ile Thr Pro Leu Leu Gln Pro Gly Asp Val Leu
 1 5 10 15
 Val Asp Gly Gly Asn Ala Tyr Phe Gly Asp Thr Arg Arg Arg Glu Glu
 20 25 30
 Glu Ile Arg Pro Thr Gly Ile His Tyr Val Gly Thr Gly Ile Ser Gly
 35 40 45
 Gly Gly Val Gly Ala Leu Arg Val Pro Ser Ile Met Pro Gly Gly Val
 50 55 60
 Lys Glu Ser Tyr Glu Ile Ile Gly Pro Val Leu Glu Lys Ile Ser Ala
 65 70 75 80
 His Val Asp Gly Glu Pro Cys Cys Ala Trp Met Gly Thr Asp Gly Ala
 85 90 95
 Gly His Phe Val Lys Met Val His Asn Gly Ile Glu Tyr Ala Asp Met
 100 105 110
 Gln Phe Ile Gly Glu Ala Pro Phe Leu Phe Ala Xaa Pro Ala Gly Leu
 115 120 125
 Thr Asn Ala Glu Ala Ala Asp Ala Phe Glu Ser Trp Asn His Gly Asp
 130 135 140
 Leu Asn Ser Tyr Leu Val Glu Ile Thr Ser Arg Val Leu Arg Ala Lys
 145 150 155 160
 Asp

<210> 399
 <211> 314
 <212> DNA
 <213> Homo sapiens

<400> 399
 nngggaatga agaccaccca gcccttcctt tcctcaaacc ttctccaggc ttctgtgcat
 60
 gggtcatcca cccatccact cattcaccca tctatccatc cactcatcca cccatccagt
 120
 cattcactca ttgtccatc cactcatgta cccatccact cattcgccca tttatccatc
 180
 cactcaacca tccactcatc cacccatcca nctcatcatc cgtccagtca cccatctatc
 240
 cacccatgta tccatccact catccaccca tccactcatc tgtccatcca cttatccacc
 300

catctactca ccca

314

<210> 400

<211> 104

<212> PRT

<213> Homo sapiens

<400> 400

Xaa Gly Met Lys Thr Thr Gln Pro Phe Leu Ser Ser Asn Leu Leu Gln
 1 5 10 15
 Ala Ser Val His Gly Ser Ser Thr His Pro Leu Ile His Pro Ser Ile
 20 25 30
 His Pro Leu Ile His Pro Ser Ser His Ser Leu Ile Cys Pro Ser Thr
 35 40 45
 His Val Pro Ile His Ser Phe Ala His Leu Ser Ile His Ser Thr Ile
 50 55 60
 His Ser Ser Thr His Pro Xaa His His Pro Ser Ser His Pro Ser Ile
 65 70 75 80
 His Pro Cys Ile His Pro Leu Ile His Pro Ser Thr His Leu Ser Ile
 85 90 95
 His Leu Ser Thr His Leu Leu Thr
 100

<210> 401

<211> 2165

<212> DNA

<213> Homo sapiens

<400> 401

gagaaaatgg aactacctgt atataaatta ggtgagcaaa cagtgatata ggtagtttta
 60
 agaagcaaat atatacagtc aatttaacag tgtttacttc tctggattgt ttaatgggtg
 120
 caaaatgaaa gatctattga agtttcacta tacattgcat tgattgaacc ttggagagtt
 180
 ttatgaaaaa gaggggcatc ccttgccatc tgtttgccag tcttccttgc cccttccttt
 240
 gaaatgcctg cctctttttt gccagattg tttcctgacc atccgaactc agatgggggc
 300
 ctctaagtgc ttctggata ttcacaaatc cttcacaag gccacgtgc gaagtgaatg
 360
 atctggaggt gcctgggcat ctgtgttgga agggagtcaa gactcaccag ccagtcagtt
 420
 tgtgggtac agttgtccca caaaaatcag gcatgttcac ctcccctctg ggcccctaca
 480
 gctgggactg atcatagcct cagattagaa gaaatactga cttctaactc tataagccag
 540
 cactcctggg taaggagtga agctctgttg gccatgccgc tttggactgc tgggcagagc
 600
 tgagcctaca gttttgtact ggggtgcacg gatgacagct gggaagatgg aaaggcagct
 660
 tgaggattta tagcagctaa agggtaaagt ctgttatgca aaagggtccc atatgaactt
 720

cctacaggtg tagccgcagc caagtgtctg tacagctgct gagaatttgt cggatgatga
780
aaaattcctc tttgcatcac aagcgagtgg aaagccaggg gctgcatgag tggagaaagc
840
acagtctggt ttttcaagta ctgcagagaa tgagaatacc cagccgggag cctggagttg
900
aggcccaggt tacacaggct cccggaatac agacctggga agatagggga ggagagggga
960
agcttgtggc cttttgatcc gccccggaa tgcccaccgt gcgctgcttt gctgccttca
1020
tctcctgctc agaggccttc tccttcccag agacctcctt ggatgggtct aaggagagaca
1080
ctgcccgggc ctttttccct gcaatcacia ggtccaaatc ctccaggctg cgcttgatcg
1140
gccgcgcgc cccaatgttc tacgggtca ttttccgggt caggattggg tggaccatgc
1200
cttccatctt cctgaaatcc tccagtctca catggtgagg ttttctgat cttgaaagcg
1260
attcagggtg ttttttaggg cctgacatgg tcatgggtga taccgacag gctttggggt
1320
gacagtctcg actctggctg cctaagacct ggaactggga gatgcctttg ctctcctggg
1380
gccctgtggt ggaatgagcc agggccagga ccttgccggg aggtttgtgc gggttcttgg
1440
gaaggctcag atctgtaggc tgatcatccg taggggcttc tgctgccgcc gactttttgt
1500
cttgacaggt cagggacgtg agataattta catggagctt ttcttggtgt ctgtgggaag
1560
gaaaagaact gttttccgat tcctgtaca tgccttga agggatattg gatgtctgtt
1620
cattatgaag atggtgctcg gtgtgtctgt agaggctatg gagatgaggg gacgagtaga
1680
agtcagccag gaagctaggc atgtgggaat gggggagggc ctttttctct aagagtttat
1740
ccttgccctc ctgaatttct tgcttcagga cgtaggagtc agcaagggg ttaagtgat
1800
gcttgagaa gctgcagcgg tgggatctg atcgactcag tttctcatgc ttaaagatgt
1860
cattgatggt ctttctctct tccgagggt tgcttctgaa actctggacg tgctgaatca
1920
ctgatggcg gctgaccgcc atatggtcag tgctttggcc atggtgggtc tgggacaaac
1980
tggaacacaa gtcatcccta gcaatcagtt tcttttgct gatcaaagg ggtggggagc
2040
cataagggtg gctgctggag aggtggccc cactcacttg ggacaaaagc ttttcttgg
2100
ccagtggga catcatgcct gggttgcccc tagagtagag caggggcgtg taattaagtc
2160
catgg
2165

<210> 402

<211> 87

<212> PRT

<213> Homo sapiens

<400> 402

```

Glu Tyr Pro Ala Gly Ser Leu Glu Leu Arg Pro Glu Leu His Arg Leu
 1             5             10             15
Pro Glu Tyr Arg Pro Gly Lys Ile Gly Glu Glu Arg Gly Ser Leu Trp
      20             25             30
Pro Phe Asp Pro Pro Pro Glu Cys Pro Pro Cys Ala Ala Leu Leu Pro
      35             40             45
Ser Ser Pro Ala Gln Arg Pro Ser Pro Ser Gln Arg Pro Pro Trp Met
      50             55             60
Gly Leu Arg Glu Thr Leu Pro Gly Pro Phe Ser Leu Gln Ser Gln Gly
65             70             75             80
Pro Asn Pro Pro Gly Cys Ala
      85

```

<210> 403

<211> 369

<212> DNA

<213> Homo sapiens

<400> 403

```

cccatgggtg tgtcccagga cggcgtcatg aagcgtcagg taaatgacaa ggaaacggtc
60
gcgcacttgt tcgaatacac gacgcaagtg tctgtcgact cgacgccgca actcgtccag
120
ccttcgcccc cgtcgcacga caacctcgtg cctgtccaga tgatcttttg cttcaagcag
180
cgcaacgcga aaaagatcaa tagccaccgc tgggtatttc atgcactggg ccgcatgcta
240
cagcccgaca tggtcgtcctt ggtggacgtc ggcacgaagc ccggccacct cgcctatac
300
catctatggc aggcattcta tcaccgacct accttgggcg gtgcttgccg cgaaattcat
360
gctatgatc
369

```

<210> 404

<211> 123

<212> PRT

<213> Homo sapiens

<400> 404

```

Pro Met Gly Val Ser Gln Asp Gly Val Met Lys Arg Gln Val Asn Asp
 1             5             10             15
Lys Glu Thr Val Ala His Leu Phe Glu Tyr Thr Thr Gln Val Ser Val
      20             25             30
Asp Ser Thr Pro Gln Leu Val Gln Pro Ser Pro Thr Ser His Asp Asn
      35             40             45
Leu Val Pro Val Gln Met Ile Phe Cys Phe Lys Gln Arg Asn Ala Lys
      50             55             60
Lys Ile Asn Ser His Arg Trp Val Phe His Ala Leu Gly Arg Met Leu
65             70             75             80
Gln Pro Asp Met Val Val Leu Val Asp Val Gly Thr Lys Pro Gly His

```

```

      85              90              95
Leu Ala Leu Tyr His Leu Trp Gln Ala Phe Tyr His Arg Pro Thr Leu
      100              105              110
Gly Gly Ala Cys Gly Glu Ile His Ala Met Ile
      115              120

```

<210> 405
 <211> 840
 <212> DNA
 <213> Homo sapiens

```

<400> 405
gaattcccg gcaccagctc gaagctggag cacttttgt ctatcctgct gaagtgttc
60
gactcgccct ggaccacgag ggcctgtcg gagacagtgg tggaggagag cgacccaag
120
ccggccttca gcaagatgaa tgggtccatg gacaaaaagt catcgaccgt cagtgaggac
180
gtggaggcca ccgtgcccat gctgcagcgg accaagtcac ggatcgagca gggatcgtg
240
gaccgctcag agacgggctg gctggacaag aaggaggggg agcaagccaa ggcgctgtt
300
gagaaggtga agaagttccg gacccatgtg gaggaggggg acattgtgta ccgctctac
360
atgcggcaga ccatcatcaa ggtgatcaag ttcacctca tcattgcta caccgtctac
420
tacgtgcaca acatcaagtt cgacgtggac tgcaccgtgg acattgagag cctgacgggc
480
taccgcacct accgctgtgc ccacccctg gccacactct tcaagatcct ggcgtccttc
540
tacatcagcc tagtcatctt ctacggctc atctgcatgt atacactgtg gtggatgcta
600
cggcgctccc tcaagaagta ctggttgag tcgatccgtg aggagagcag ctacagcgac
660
atccccgagc tcaagaacga ctgcgcttc atgctgcacc tcattgacca atacgaccg
720
ctctactcca agcgcttcgc cgtcttctg tcggaggtga gtgagaacaa gctgcggcag
780
ctgaacctca acaacgagtg gacgctggac aagctccggt acggagagaa gacaacgcgt
840

```

<210> 406
 <211> 91
 <212> PRT
 <213> Homo sapiens

```

<400> 406
Leu Ile Cys Met Tyr Thr Leu Trp Trp Met Leu Arg Arg Ser Leu Lys
1              5              10              15
Lys Tyr Ser Phe Glu Ser Ile Arg Glu Glu Ser Ser Tyr Ser Asp Ile
20              25              30
Pro Asp Val Lys Asn Asp Phe Ala Phe Met Leu His Leu Ile Asp Gln
35              40              45
Tyr Asp Pro Leu Tyr Ser Lys Arg Phe Ala Val Phe Leu Ser Glu Val

```


50 55 60
 Ser Glu Asn Lys Leu Arg Gln Leu Asn Leu Asn Glu Trp Thr Leu
 65 70 75 80
 Asp Lys Leu Arg Tyr Gly Glu Lys Thr Thr Arg
 85 90

<210> 407
 <211> 535
 <212> DNA
 <213> Homo sapiens

<400> 407
 gcctattgta ccagctctcc agggctgggg acttgctaga gcagggttcc cagtgcctccc
 60
 aggtcttact ttgctctgcc tggctctcagg gtgtagggga tggagagctg gacttccagc
 120
 ctgtctcttg gctgtctagg ggccaggggc tcgggacaca gagctcctgg aggccgagca
 180
 caagccttgg gcagaggtga ggcagagctc tgactgtttc attcgactac gttgccaagg
 240
 agatgctcgc tcggagtggg tgcctctggc ctgggattcc aaaccaagct gccttctctg
 300
 atgtggcctt agtgcctctg gcggatgtac cttggctctg cctggaccct ctctctcttc
 360
 caggcctctg tcccaccagg atgatgccta tccagagctc attgtcctct cccacttctt
 420
 ccccgagctt cccattccgt gtctctctgg agggcccatc atcatcctgg tggaggtgtt
 480
 gcactgagga ccacagcagc cctcgcatc ccacgggcaa aggggtatgt gtagg
 535

<210> 408
 <211> 97
 <212> PRT
 <213> Homo sapiens

<400> 408
 Met Leu Ala Arg Ser Gly Cys Ser Gly Ser Gly Ile Pro Asn Gln Ala
 1 5 10 15
 Ala Phe Ser Asp Val Ala Leu Val Leu Trp Ala Asp Val Pro Trp Leu
 20 25 30
 Cys Leu Asp Pro Leu Ser Leu Pro Gly Leu Cys Pro Thr Arg Met Met
 35 40 45
 Pro Ile Gln Ser Ser Leu Ser Ser Pro Thr Ser Ser Pro Ser Phe Pro
 50 55 60
 Phe Arg Val Ser Leu Glu Gly Pro Ser Ser Ser Trp Trp Arg Cys Cys
 65 70 75 80
 Thr Glu Asp His Ser Ser Pro Arg Ile Pro Thr Gly Lys Gly Val Cys
 85 90 95
 Val

<210> 409
 <211> 375

<212> DNA

<213> Homo sapiens

<400> 409

ngtgtcatgg gtgtctatac cagcgatgag gccaaagactg ccaagacttt tggatttggt
 60
 ggacttccga ttacgactaa tatttctctt gccacaact tcaatatgga tgaaatttct
 120
 gatattgtct tccgtgtcaa tgataccagt ttgacaccaa ctgtgggacc agaattagct
 180
 agaaaattga ccgaaattgc tggctctcag caaggggagt atcaggtgtc agatgcgact
 240
 gcagccttcc aagaagtgc acaattgttc ggctttataa ctacgattat tagtgccatt
 300
 gcaggaattt ccctttttgt tggagggact ggtgttatga acatcatgct ggtttcggtg
 360
 acggagcgta cgcgt
 375

<210> 410

<211> 125

<212> PRT

<213> Homo sapiens

<400> 410

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Xaa | Val | Met | Gly | Val | Tyr | Thr | Ser | Asp | Glu | Ala | Lys | Thr | Ala | Lys | Thr |
| 1 | | | 5 | | | | | | 10 | | | | 15 | | |
| Phe | Gly | Ile | Gly | Gly | Leu | Pro | Ile | Thr | Thr | Asn | Ile | Ser | Leu | Ala | Asn |
| | | | 20 | | | | | 25 | | | | | 30 | | |
| Asn | Phe | Asn | Met | Asp | Glu | Ile | Ser | Asp | Ile | Val | Phe | Arg | Val | Asn | Asp |
| | | | 35 | | | | | 40 | | | | 45 | | | |
| Thr | Ser | Leu | Thr | Pro | Thr | Val | Gly | Pro | Glu | Leu | Ala | Arg | Lys | Leu | Thr |
| | | | 50 | | | | 55 | | | | 60 | | | | |
| Glu | Ile | Ala | Gly | Leu | Gln | Gln | Gly | Glu | Tyr | Gln | Val | Ser | Asp | Ala | Thr |
| 65 | | | | | 70 | | | | 75 | | | | | 80 | |
| Ala | Ala | Phe | Gln | Glu | Val | Gln | Gln | Leu | Phe | Gly | Phe | Ile | Thr | Thr | Ile |
| | | | 85 | | | | | 90 | | | | | 95 | | |
| Ile | Ser | Ala | Ile | Ala | Gly | Ile | Ser | Leu | Phe | Val | Gly | Gly | Thr | Gly | Val |
| | | | 100 | | | | | 105 | | | | | 110 | | |
| Met | Asn | Ile | Met | Leu | Val | Ser | Val | Thr | Glu | Arg | Thr | Arg | | | |
| | | | 115 | | | | | 120 | | | | 125 | | | |

<210> 411

<211> 409

<212> DNA

<213> Homo sapiens

<400> 411

ccacatactt caccctcttc acccctcca cctactccac cacctggcag tcgccatcga
 60
 ggatgggacg caactccacg tccacatgct ccggaccacg cggcgtgtgg tggatgtgca
 120
 gcacgcggtc ggggccctt gagctcgaag gcgcggcgca tcgggcagtg ctgcgcggcc
 180

tggtcgcagg gcacgtcgta ctggtgagac acgcggaagc acttgtggcc gatgtaggcg
 240
 cgatcggctg tcccgaactg gcgctgatag gccgtgtaca caacacaaac tgttgtactc
 300
 ccggtccacc acgatcatgg gctgggactc gtgttccagg tggggggcca gggcttgggc
 360
 ctgcggtgag cgcgtggggg ggatggggca tagcgtcggg gaggaggtg
 409

<210> 412
 <211> 119
 <212> PRT
 <213> Homo sapiens

<400> 412
 Met Pro His Pro Pro His Ala Leu Thr Ala Gly Pro Ser Pro Gly Pro
 1 5 10 15
 Pro Pro Gly Thr Arg Val Pro Ala His Asp Arg Gly Gly Pro Gly Val
 20 25 30
 Gln Gln Phe Val Leu Cys Thr Arg Pro Ile Ser Ala Ser Ser Gly Gln
 35 40 45
 Pro Ile Ala Pro Thr Ser Ala Thr Ser Ala Ser Ala Ser Arg Thr Ser
 50 55 60
 Thr Thr Cys Pro Ala Thr Arg Pro Ala Ser Thr Ala Arg Cys Ala Ala
 65 70 75 80
 Pro Ser Ser Ser Arg Gly Pro Asp Arg Val Leu His Ile His His Thr
 85 90 95
 Pro Arg Gly Pro Glu His Val Asp Val Glu Leu Arg Pro Ile Leu Asp
 100 105 110
 Gly Asp Cys Gln Val Val Glu
 115

<210> 413
 <211> 357
 <212> DNA
 <213> Homo sapiens

<400> 413
 ccgggcatcc caccaccggg tgtcatgaac caagtagtgg cccctatggt agggactcca
 60
 gcaccgggtg gaagtccata tggacaacag gtgggagttt tggggcctcc agggcagcag
 120
 gcaccacctc catatcccgg cccacatcca gctggacccc ctgtcataca gcagccaaca
 180
 acacccatgt ttgtagctcc cccccaaag acccagcggc ttcttcactc agaggcctac
 240
 ctgaaatata ttgaaggact cagtgcggag tccaacagca ttagcaagtg ggatcagaca
 300
 ctggcagctc ggagacgcga cgtccatttg tcgaaagaac aggagagccg cctaccc
 357

<210> 414
 <211> 119
 <212> PRT

<213> Homo sapiens

<400> 414

```

Pro Gly Ile Pro Pro Pro Gly Val Met Asn Gln Val Val Ala Pro Met
 1           5           10           15
Val Gly Thr Pro Ala Pro Gly Gly Ser Pro Tyr Gly Gln Gln Val Gly
      20           25           30
Val Leu Gly Pro Pro Gly Gln Gln Ala Pro Pro Pro Tyr Pro Gly Pro
      35           40           45
His Pro Ala Gly Pro Pro Val Ile Gln Gln Pro Thr Thr Pro Met Phe
      50           55           60
Val Ala Pro Pro Pro Lys Thr Gln Arg Leu Leu His Ser Glu Ala Tyr
      65           70           75           80
Leu Lys Tyr Ile Glu Gly Leu Ser Ala Glu Ser Asn Ser Ile Ser Lys
      85           90           95
Trp Asp Gln Thr Leu Ala Ala Arg Arg Asp Val His Leu Ser Lys
      100          105          110
Glu Gln Glu Ser Arg Leu Pro
      115

```

<210> 415

<211> 332

<212> DNA

<213> Homo sapiens

<400> 415

```

tctagagcca acttggttat cgtaatgaat agagagacta catctatatc aattattacg
60
ctctatagta atcatgaagc ttgggttata tgtatgacaa aaattgcaga aaaatcgaaa
120
caagaatatg gcgacttact aaaagaaaaa gaccatttac aagatatgga acagcttgag
180
atgactatcg tctcgatcca tacgccgtat ccgtccattg tcagaattca aggaaaaatc
240
aacacattac agccagagct ttggcaagct cccaatttag caattcggtt aattgtgagc
300
aatccgccag agggacaacc catctcacgc gt
332

```

<210> 416

<211> 102

<212> PRT

<213> Homo sapiens

<400> 416

```

Met Asn Arg Glu Thr Thr Ser Ile Ser Ile Ile Thr Leu Tyr Ser Asn
 1           5           10           15
His Glu Ala Trp Val Ile Cys Met Thr Lys Ile Ala Glu Lys Ser Lys
      20           25           30
Gln Glu Tyr Gly Asp Leu Leu Lys Lys Asp His Leu Gln Asp Met
      35           40           45
Glu Gln Leu Glu Met Thr Ile Val Ser Ile His Thr Pro Tyr Pro Ser
      50           55           60
Ile Val Arg Ile Gln Gly Lys Ile Asn Thr Leu Gln Pro Glu Leu Trp

```

```

<400> 418
Glu Phe Leu Ala Val Ser Glu Val Gly Glu Asp Thr Phe Val Arg Ser
 1             5             10             15
Thr Glu Gly Asp Tyr Ala Ala Asn Val Glu Ala Val Val Thr Pro Ala
          20             25             30
Pro Ala Glu Lys Asp Ile Glu Gly Gln Pro Glu Ala Gln Glu His Asp
          35             40             45
Thr Pro Gly Thr Glu Thr Ile Glu Lys Leu Val Glu Trp Ala Gln Gly
          50             55             60
Ala Gly Ile Thr Val Asn Pro Arg Val Val Cys Tyr Tyr Thr Leu Lys
65             70             75             80
Cys Met Met Ile Lys Leu His His Pro Ala Ala Glu Ser Glu Glu Arg
          85             90             95
Glu Ser Glu Leu Ala Ala Val Leu Ile Pro Gly Asp Arg Glu Leu Asp
          100             105             110
Glu Lys Arg Leu Glu Ala Ala Leu Glu Pro Val Glu Phe Glu Leu Ala
          115             120             125
Gly Asp Lys Asp Phe Ala Asp Asn Asp Phe Leu Val Lys Gly Tyr Val

```

| | | |
|---|-----|-----|
| 130 | 135 | 140 |
| Gly Pro Arg Ala Leu Asn Ala Asn Gly Ile Lys Val Leu Ala Asp Pro | | |
| 145 | 150 | 155 |
| Arg | | 160 |

<210> 419
 <211> 797
 <212> DNA
 <213> Homo sapiens

<400> 419
 atttcacccc aggaaaacca gtaaggacca atgattaagc ccaagggttg gtaccgagtt
 60
 cggatccata agtaccggcc gccaggggtg ctggaatttg ggctcccccc ggtgaaaaata
 120
 tccatgcagc cgcgttgtct taggtagaaa agggagactg ggggtgggtg ggctgagctc
 180
 aagcccctgc ctacatactt tagtagtaac gactcccgat ctgcatccaa cacatttacc
 240
 gaacttctag taagcgcccc ccgctgcaag cgaaagcact cccctgccaa gaaacagatc
 300
 ttttcactt aaaattccca aactcagacc ttccactttt tactgaacaa aaagcgtgta
 360
 catgatctga agggttgaca tgacattttc taaattgggc gaatcaggaa gaggttgatg
 420
 aaaatccttg acgtttttctg gggataggac atttgtgtgt gataacgttc ttaagtcgaa
 480
 tttcagtgtg gcagtgcacg cagattcttc attggtgtta gtgtatttcc atacggtatg
 540
 tattagtaca agaaatagtg ttccctttga cactcgaacc caaggagtgg tccgaggctt
 600
 tttgaggcaa cgtaggatca atgtctctga agcagatttg gtgaaggatg caggtctcat
 660
 aatttacaga gcaatcacag ccttctttga aacggagaaa ttagattcta tgaaattttg
 720
 tcagtgcaga tagatatgat gtggagaaac ggggaaaatt gagtacaaaa agatgaggct
 780
 tgaatgatgg ctggcca
 797

<210> 420
 <211> 106
 <212> PRT
 <213> Homo sapiens

<400> 420
 Met Arg Pro Ala Ser Phe Thr Lys Ser Ala Ser Glu Thr Leu Ile Leu
 1 5 10 15
 Arg Cys Leu Lys Lys Pro Arg Thr Thr Pro Trp Val Arg Val Ser Lys
 20 25 30
 Gly Thr Leu Phe Leu Val Leu Ile His Thr Val Trp Lys Tyr Thr Asn
 35 40 45
 Thr Asn Glu Glu Ser Ala Cys Thr Ala Thr Leu Lys Phe Asp Leu Arg

```

      50              55              60
Thr Leu Ser His Thr Asn Val Leu Ser Pro Glu Asn Val Lys Asp Phe
65              70              75              80
His Gln Pro Leu Pro Asp Ser Pro Asn Leu Glu Asn Val Met Ser Thr
      85              90              95
Leu Gln Ile Met Tyr Thr Leu Phe Val Gln
      100              105

```

<210> 421
 <211> 406
 <212> DNA
 <213> Homo sapiens

```

<400> 421
ggatccacca tgatggagcc caccacacca tcctcagtc accctgctgca gcttctccat
60
aaccacaacac aggtcaatct tgtctcccta aacacaccat gtgctctcat gctgccatgg
120
tttgctggg gccctctcta cctcctctgc tttctggaga acccttgac tctcccaag
180
ccttcaagtt ggaaagtga cagtcagcat atgtctctag ctcagccctt actgcgtgga
240
ttcatgaaga ttggttcaact gtcagccctt gaccagaacg tgtgttttag gaaagcagga
300
accaagtctt accaatgtct gtagtccag cctccaccct ggcatacagt aggtgctcat
360
tgaatgtggg agggaaagag gagacacatg gaagggaatg tcattc
406

```

<210> 422
 <211> 104
 <212> PRT
 <213> Homo sapiens

```

<400> 422
Met Met Glu Pro Thr His Pro Ser Ser Val His Leu Leu Gln Leu Leu
1              5              10              15
His Asn Pro Thr Gln Val Asn Leu Val Ser Leu Asn Thr Pro Cys Ala
      20              25              30
Leu Met Leu Pro Trp Phe Ala Trp Gly Pro Leu Tyr Leu Leu Cys Phe
      35              40              45
Leu Glu Asn Pro Cys Thr Pro Pro Lys Pro Ser Ser Trp Lys Val Asn
      50              55              60
Ser Gln His Met Ser Leu Ala Gln Pro Leu Leu Arg Gly Phe Met Lys
65              70              75              80
Ile Gly Ser Leu Ser Ala Pro Asp Gln Asn Val Cys Phe Arg Lys Ala
      85              90              95
Gly Thr Lys Ser Tyr Gln Cys Leu
      100

```

<210> 423
 <211> 628
 <212> DNA
 <213> Homo sapiens

<400> 423

ngccacccta cgcctcgcct gcaatggcaa cttcagatcc ccggtggcac cgtagtctta
 60
 gagccaccgg ttctgagcgg ggaggacgac ggggttgggg cgagaggaagg agagggagaa
 120
 ggagatgggg atttgctgac gcagacccaa gccc aaacgc cgactccagc acccgcttgg
 180
 ccggcgcccc cagccacacc gcgcttcttg gccctcgcaa atggctccct gttggtgccc
 240
 ctcttgagtg ccaaggaggc gggcgtctac acttgccgtg cacacaatga gctgggcgcc
 300
 aactctacgt caatacgcgt ggcggtggca gcaaccgggc ccccaaaaca cgcgcctggc
 360
 gccggggggag aaccgcacgg acaggccccg acctctgagc gcaagtccac agccaagggc
 420
 cggggcaaca gcgtcctgcc ttccaaaccc gagggcaaaa tcaaaggcca aggcctggcc
 480
 aaggtcagca ttctcgggga gaccgagacg gagccggagg aggacacaag tgagggagag
 540
 gaggccgaag accagatcct cgcggaccgc gcggaggagc agcgtgtgg caacggggac
 600
 ccctctcggg acgtttctaa ccacgcgt
 628

<210> 424

<211> 209

<212> PRT

<213> Homo sapiens

<400> 424

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Xaa | His | Pro | Thr | Pro | Arg | Leu | Gln | Trp | Gln | Leu | Gln | Ile | Pro | Gly | Gly |
| 1 | | | | 5 | | | | 10 | | | | | | 15 | |
| Thr | Val | Val | Leu | Glu | Pro | Pro | Val | Leu | Ser | Gly | Glu | Asp | Asp | Gly | Val |
| | | | 20 | | | | | 25 | | | | | 30 | | |
| Gly | Ala | Glu | Gly | Glu | Gly | Gly | Glu | Gly | Asp | Gly | Asp | Leu | Leu | Thr | Gln |
| | | 35 | | | | 40 | | | | | | 45 | | | |
| Thr | Gln | Ala | Gln | Thr | Pro | Thr | Pro | Ala | Pro | Ala | Trp | Pro | Ala | Pro | Pro |
| | | 50 | | | | 55 | | | | | 60 | | | | |
| Ala | Thr | Pro | Arg | Phe | Leu | Ala | Leu | Ala | Asn | Gly | Ser | Leu | Leu | Val | Pro |
| 65 | | | | 70 | | | | | 75 | | | | | 80 | |
| Leu | Leu | Ser | Ala | Lys | Glu | Ala | Gly | Val | Tyr | Thr | Cys | Arg | Ala | His | Asn |
| | | | 85 | | | | | 90 | | | | | 95 | | |
| Glu | Leu | Gly | Ala | Asn | Ser | Thr | Ser | Ile | Arg | Val | Ala | Val | Ala | Ala | Thr |
| | | 100 | | | | | 105 | | | | | | 110 | | |
| Gly | Pro | Pro | Lys | His | Ala | Pro | Gly | Ala | Gly | Gly | Glu | Pro | Asp | Gly | Gln |
| | | 115 | | | | 120 | | | | | 125 | | | | |
| Ala | Pro | Thr | Ser | Glu | Arg | Lys | Ser | Thr | Ala | Lys | Gly | Arg | Gly | Asn | Ser |
| | | 130 | | | | 135 | | | | | 140 | | | | |
| Val | Leu | Pro | Ser | Lys | Pro | Glu | Gly | Lys | Ile | Lys | Gly | Gln | Gly | Leu | Ala |
| 145 | | | | 150 | | | | | 155 | | | | | 160 | |
| Lys | Val | Ser | Ile | Leu | Gly | Glu | Thr | Glu | Thr | Glu | Pro | Glu | Glu | Asp | Thr |
| | | | 165 | | | | | 170 | | | | | 175 | | |
| Ser | Glu | Gly | Glu | Glu | Ala | Glu | Asp | Gln | Ile | Leu | Ala | Asp | Pro | Ala | Glu |

180 185 190
 Glu Gln Arg Cys Gly Asn Gly Asp Pro Ser Arg Tyr Val Ser Asn His
 195 200 205
 Ala

<210> 425
 <211> 471
 <212> DNA
 <213> Homo sapiens

<400> 425
 ccggccgctcg aagactttga ggacgatgta gctcgacgag cagcggttacg agccctggag
 60
 tacgtggatt tgaccccagg cactnaagtg cgcgtcatcg ccattgacac cgtgttccta
 120
 ggatcgtgca cgaatggccg tgaggactta cggctggctg ctgagggttc caaaggacga
 180
 catatcgag cgggcacccg gatgctcgtc gccctggat ctgctcgtgt ccgtctgcag
 240
 gctatggagg aaggcctcga cgagatcggg tcccggtttg ctgacatctt tcgcaataac
 300
 tctgccaaca atggcttggt actggctcag gttgacccc aggtcgtcga agagttgtgg
 360
 gactttgccg agcagcatcc tggtagcag ctcaccgtct ccctcgagaa tcggacgatc
 420
 aaccttcggt gtcgcacgac ctaccgttc catattgatg acgtcacgag t
 471

<210> 426
 <211> 157
 <212> PRT
 <213> Homo sapiens

<400> 426
 Pro Ala Val Glu Asp Phe Glu Asp Asp Val Ala Arg Ser Ala Ala Leu
 1 5 10 15
 Arg Ala Leu Glu Tyr Val Asp Leu Thr Pro Gly Thr Xaa Val Arg Val
 20 25 30
 Ile Ala Ile Asp Thr Val Phe Leu Gly Ser Cys Thr Asn Gly Arg Glu
 35 40 45
 Asp Leu Arg Leu Ala Ala Glu Val Pro Lys Gly Arg His Ile Ala Ala
 50 55 60
 Gly Thr Arg Met Leu Val Ala Pro Gly Ser Ala Arg Val Arg Leu Gln
 65 70 75 80
 Ala Met Glu Glu Gly Leu Asp Glu Ile Gly Ser Arg Phe Ala Asp Ile
 85 90 95
 Phe Arg Asn Asn Ser Ala Asn Asn Gly Leu Leu Leu Ala Gln Val Asp
 100 105 110
 Pro Glu Val Val Glu Glu Leu Trp Asp Phe Ala Glu Gln His Pro Gly
 115 120 125
 Glu Gln Leu Thr Val Ser Leu Glu Asn Arg Thr Ile Asn Leu Pro Gly
 130 135 140
 Arg Thr Thr Tyr Pro Phe His Ile Asp Asp Val Thr Arg

145

150

155

<210> 427

<211> 546

<212> DNA

<213> Homo sapiens

<400> 427

ctagcggtag tagaaggat gcagtttgat cgcggctact tgtctccgta tttcatcaac
 60
 aatcaagaaa caatgaatgc agagctagaa aaccatttta ttcttcttgt tgataagaaa
 120
 atttctaata tccgtgactt gctaccaatt ttggaagggt ttgctaaagc atcgcgcccc
 180
 ttgttgatca ttgcggaaga cgttgaaggc gaagcgttgg caaccttggg tgtaaacact
 240
 atgcgcggca tcgtaaaaagt agcggcagcg aaagcggcag gttttggtga tcgccgtaaa
 300
 gcaatgcttc aagacattgc tgtgctaacg ggttcaactg ttatttcaga agaaattggc
 360
 attaagcttg aagaagcgac aattgaacag ttgggtacag cgaagcgcgt tacattgaca
 420
 aaagaaagta caacgattgt tgatggtgcg ggtgttgacg ctaatattac tggtcgtggt
 480
 gagcaaattc gtgcagaaat tgctaactct tcttctggct acgataaaga gaaattgcaa
 540
 gaacgc
 546

<210> 428

<211> 182

<212> PRT

<213> Homo sapiens

<400> 428

Leu Ala Val Val Glu Gly Met Gln Phe Asp Arg Gly Tyr Leu Ser Pro
 1 5 10 15
 Tyr Phe Ile Asn Asn Gln Glu Thr Met Asn Ala Glu Leu Glu Asn Pro
 20 25 30
 Phe Ile Leu Leu Val Asp Lys Lys Ile Ser Asn Ile Arg Asp Leu Leu
 35 40 45
 Pro Ile Leu Glu Gly Val Ala Lys Ala Ser Arg Pro Leu Leu Ile Ile
 50 55 60
 Ala Glu Asp Val Glu Gly Glu Ala Leu Ala Thr Leu Val Val Asn Thr
 65 70 75 80
 Met Arg Gly Ile Val Lys Val Ala Ala Ala Lys Ala Pro Gly Phe Gly
 85 90 95
 Asp Arg Arg Lys Ala Met Leu Gln Asp Ile Ala Val Leu Thr Gly Ser
 100 105 110
 Thr Val Ile Ser Glu Glu Ile Gly Ile Lys Leu Glu Glu Ala Thr Ile
 115 120 125
 Glu Gln Leu Gly Thr Ala Lys Arg Val Thr Leu Thr Lys Glu Ser Thr
 130 135 140
 Thr Ile Val Asp Gly Ala Gly Val Ala Ala Asn Ile Thr Gly Arg Val

145 150 155 160
 Glu Gln Ile Arg Ala Glu Ile Ala Asn Ser Ser Ser Gly Tyr Asp Lys
 165 170 175
 Glu Lys Leu Gln Glu Arg
 180

<210> 429
 <211> 425
 <212> DNA
 <213> Homo sapiens

<400> 429
 gctagcagcc cttacaggag acgggctaataataatgcag cagtggctcc gacaacttgc
 60
 ccgttgacgc cggtcacgga tccatttgct tttagtagac aggcgctcca aagtacacca
 120
 ctgggcagtt cgtccaaaag cagtccacct gtcttgcaag gccagcccc cgcagggttt
 180
 tctcaacacc ccggtttgct tgtgccttac acacaatgca aaaaatagct ctcaggggacc
 240
 ctgtgagccc ctgcctggac ctctgacaca gccagagca catgccagtc cgttttctgg
 300
 tgcattgaca ccttcagcac ctctggggcc tgagatgaac aggagtgcag aggtcgggtcc
 360
 cagttcagag cctgaagttc agactctgcc atatcttcct cactacattc caggagtgga
 420
 tctctg
 425

<210> 430
 <211> 130
 <212> PRT
 <213> Homo sapiens

<400> 430
 Met Gln Gln Trp Leu Arg Gln Leu Ala Arg Cys Ser Arg Ser Arg Ile
 1 5 10 15
 His Leu Leu Leu Val Asp Arg Arg Ser Lys Val His His Trp Ala Val
 20 25 30
 Arg Pro Lys Ala Val His Leu Ser Cys Lys Ala Gln Pro Pro Gln Gly
 35 40 45
 Phe Leu Asn Thr Pro Val Cys Leu Cys Leu Thr His Asn Ala Lys Asn
 50 55 60
 Ser Ser Gln Gly Pro Cys Glu Pro Leu Pro Gly Pro Leu Thr Gln Pro
 65 70 75 80
 Arg Ala His Ala Ser Pro Phe Ser Gly Ala Leu Thr Pro Ser Ala Pro
 85 90 95
 Pro Gly Pro Glu Met Asn Arg Ser Ala Glu Val Gly Pro Ser Ser Glu
 100 105 110
 Pro Glu Val Gln Thr Leu Pro Tyr Leu Pro His Tyr Ile Pro Gly Val
 115 120 125
 Asp Pro
 130

<210> 431
 <211> 192
 <212> DNA
 <213> Homo sapiens

<400> 431
 ctageccatcc accagcgtac acacacggga gagaggccct acactggcct cgggtgcaac
 60
 cgccgcttcc gccagcgcac ggccctcgtc atccaccagc gcatccacac gggcgagaag
 120
 cctnaccogt gcccgactg cgagcggcgc ttctcctcct cctctcgcct ggtcagtcac
 180
 cggcgtgtgc ac
 192

<210> 432
 <211> 64
 <212> PRT
 <213> Homo sapiens

<400> 432
 Leu Ala Ile His Gln Arg Thr His Thr Gly Glu Arg Pro Tyr Thr Gly
 1 5 10 15
 Leu Gly Cys Asn Arg Arg Phe Arg Gln Arg Thr Ala Leu Val Ile His
 20 25 30
 Gln Arg Ile His Thr Gly Glu Lys Pro Xaa Pro Cys Pro Asp Cys Glu
 35 40 45
 Arg Arg Phe Ser Ser Ser Ser Arg Leu Val Ser His Arg Arg Val His
 50 55 60

<210> 433
 <211> 635
 <212> DNA
 <213> Homo sapiens

<400> 433
 nngccggcgg ctgcgttggg atacgacgtc gctgcgattg ggcgtgagta tctttggtac
 60
 ctcattggagg agcgtggcgc gtatgcggag gccgcgcgc tcatgccgct gctgctccgg
 120
 accgaccgag gcgcgtggga cacgtttgtg tgctgctacc tcgagcggca ccaaaggat
 180
 gccgatactcc cgcacattcc gacgcaggac cccagctga gtgagatggt gtacgatctc
 240
 gtgctggtgc atctgctgca gcacgatccc acgcagctgt tggcgacgct ccgcgcatgg
 300
 ccgagtcaca tctactcgaa gcaggcgggtg gctgcggcga tcggcgatca cgcacgaacc
 360
 agccgcacgc tgctcgagtg cctcgcacag ctgtacatgg ccgcacatca gcccggaag
 420
 gctctgacat actacatgcy cctgcgtgat ccatgcgtgt ttgatctcat tcgcgagtac
 480
 gatctgctga tcgatgtgca gcaccacatc ggcacgctcg tcgagctcga tcaggaatgc
 540

gcccggctcca ctgagccgcg ctccagcgcg cttatgccgc tgctcgtgcc atataccac
 600
 tcgattccca tccagcgcgc catggcgcag ctcga
 635

<210> 434
 <211> 211
 <212> PRT
 <213> Homo sapiens

<400> 434
 Xaa Pro Ala Ala Leu Gly Tyr Asp Val Ala Ala Ile Gly Arg Glu
 1 5 10 15
 Tyr Leu Trp Tyr Leu Met Glu Glu Arg Gly Ala Tyr Ala Glu Ala Ala
 20 25 30
 Ala Leu Met Pro Leu Leu Leu Arg Thr Asp Arg Gly Ala Trp Asp Thr
 35 40 45
 Phe Val Cys Cys Tyr Leu Glu Arg His Gln Arg Asp Ala Ile Leu Pro
 50 55 60
 His Ile Pro Thr Gln Asp Pro Gln Leu Ser Glu Met Val Tyr Asp Leu
 65 70 75 80
 Val Leu Val His Leu Leu Gln His Asp Pro Thr Gln Leu Leu Ala Thr
 85 90 95
 Leu Arg Ala Trp Pro Ser His Ile Tyr Ser Lys Gln Ala Val Ala Ala
 100 105 110
 Ala Ile Gly Asp His Ala Arg Thr Ser Arg Thr Leu Leu Glu Cys Leu
 115 120 125
 Ala Gln Leu Tyr Met Ala Ala His Gln Pro Gly Lys Ala Leu Thr Tyr
 130 135 140
 Tyr Met Arg Leu Arg Asp Pro Cys Val Phe Asp Leu Ile Arg Glu Tyr
 145 150 155 160
 Asp Leu Leu Ile Asp Val Gln His His Ile Gly Thr Leu Val Glu Leu
 165 170 175
 Asp Gln Glu Cys Ala Gly Ser Thr Glu Pro Arg Ser Ser Ala Leu Met
 180 185 190
 Pro Leu Leu Val Pro Tyr Thr His Ser Ile Pro Ile Gln Arg Ala Met
 195 200 205
 Ala Gln Leu
 210

<210> 435
 <211> 493
 <212> DNA
 <213> Homo sapiens

<400> 435
 nncgtacgtt cgcgtatttt ccgcgcccgg gaagctatcg ataataaagt tcaaccgctg
 60
 atccagcgtt agcaatggcg ggcacaggaa gggtagcttag gcatgcagaa agaaaagctt
 120
 tccgctctga tggatgggtga atcgttcgac agcgagctgt tgagttctct gtcgcaagat
 180
 cgaacgcttc aacaaagctg gcagggctat cacctgatac gtgacacact gcgaggtgat
 240

gtcgggcaag tgatgcatct cgacatcgcc gatcgcgtag ccgctgcact tgagaaagaa
 300
 cccgccccgc tgggtgccttc cgcggttcag gaatctcagc cgcagcctca cacctggcag
 360
 aaaaatgccgt tctgggacaa agtgcgtccc tgggagagcc agattacgca aatcgggtatg
 420
 gcggcctgcg tgtcgctggc ggtgatcgtc ggcgtgcagc agtacaacca gccttctgcg
 480
 ccatcgaacg cgt
 493

<210> 436
 <211> 130
 <212> PRT
 <213> Homo sapiens

<400> 436
 Met Gln Lys Glu Lys Leu Ser Ala Leu Met Asp Gly Glu Ser Phe Asp
 1 5 10 15
 Ser Glu Leu Leu Ser Ser Leu Ser Gln Asp Arg Thr Leu Gln Gln Ser
 20 25 30
 Trp Gln Gly Tyr His Leu Ile Arg Asp Thr Leu Arg Gly Asp Val Gly
 35 40 45
 Gln Val Met His Leu Asp Ile Ala Asp Arg Val Ala Ala Ala Leu Glu
 50 55 60
 Lys Glu Pro Ala Arg Leu Val Pro Ser Ala Val Gln Glu Ser Gln Pro
 65 70 75 80
 Gln Pro His Thr Trp Gln Lys Met Pro Phe Trp Asp Lys Val Arg Pro
 85 90 95
 Trp Ala Ser Gln Ile Thr Gln Ile Gly Met Ala Ala Cys Val Ser Leu
 100 105 110
 Ala Val Ile Val Gly Val Gln Gln Tyr Asn Gln Pro Ser Ala Pro Ser
 115 120 125
 Asn Ala
 130

<210> 437
 <211> 447
 <212> DNA
 <213> Homo sapiens

<400> 437
 ntggtaaccg gtgtccctga tatggaccct gctgtgtag agcgtaaatt atttatttta
 60
 cgtaattatg taacacgcat ctgtttggag tctgttaatg gaattaagga caacttttac
 120
 attaatacat tctcatacaa aacaatcggt tataaaggtc agttaaccac tgaacaagtg
 180
 ccacaatatt tcttagatgt acaaaatcca agtatggtaa cggcattagc gcttggtcat
 240
 tcacgtttct caacaaatac atttcctcgt tggcgtttag cacaaccatt ccgttacatc
 300
 gctcataatg gcgaaatcaa tacggttcgc ggtaatatca attggatgaa agcacgtgaa
 360

gcgttacttg aagctgaatt tttcactcgc tcagaattag atatgttaat gccaatctgt
 420
 acggatggta tgtctgactc ggcaagg
 447

<210> 438
 <211> 149
 <212> PRT
 <213> Homo sapiens

<400> 438
 Xaa Val Thr Gly Val Pro Asp Met Asp Pro Ala Val Leu Glu Arg Lys
 1 5 10 15
 Leu Phe Ile Leu Arg Asn Tyr Val Thr Arg Ile Cys Leu Glu Ser Val
 20 25 30
 Asn Gly Ile Lys Asp Asn Phe Tyr Ile Asn Thr Phe Ser Tyr Lys Thr
 35 40 45
 Ile Val Tyr Lys Gly Gln Leu Thr Thr Glu Gln Val Pro Gln Tyr Phe
 50 55 60
 Leu Asp Leu Gln Asn Pro Ser Met Val Thr Ala Leu Ala Leu Val His
 65 70 75 80
 Ser Arg Phe Ser Thr Asn Thr Phe Pro Arg Trp Arg Leu Ala Gln Pro
 85 90 95
 Phe Arg Tyr Ile Ala His Asn Gly Glu Ile Asn Thr Val Arg Gly Asn
 100 105 110
 Ile Asn Trp Met Lys Ala Arg Glu Ala Leu Leu Glu Ala Glu Phe Phe
 115 120 125
 Thr Arg Ser Glu Leu Asp Met Leu Met Pro Ile Cys Thr Asp Gly Met
 130 135 140
 Ser Asp Ser Ala Arg
 145

<210> 439
 <211> 395
 <212> DNA
 <213> Homo sapiens

<400> 439
 nacgcgtgaa gggagagtgg ggccgagccc caggaggctg tcctgcagca gctgcaccag
 60
 cttcccaggg gccggtgga cctggccacg caaagcctga cggtggagac ctgcagggcc
 120
 ctgggcaagc tgctgccgag ggagacgctg tgcacggagc tggctctgag tgactgcatg
 180
 ctcagcgagg aaggggccac actgctgctc cgaggcctgt gtgccaacac cgtgctgcgc
 240
 tttctggact taaagggcaa caaccttcgg gctgcagggg ccgaggctct gggaaaactc
 300
 ctccaacaga acaagtccat tcagagcctc acgctggagt ggaacagcct gggcacgtgg
 360
 gacgatgect tcgccacctt ctgcgggggc ctggc
 395

<210> 440

<211> 128
 <212> PRT
 <213> Homo sapiens

<400> 440
 Arg Glu Ser Gly Ala Glu Pro Gln Glu Ala Val Leu Gln Gln Leu His
 1 5 10 15
 Gln Leu Pro Arg Gly Arg Leu Asp Leu Ala Thr Gln Ser Leu Thr Val
 20 25 30
 Glu Thr Cys Arg Ala Leu Gly Lys Leu Leu Pro Arg Glu Thr Leu Cys
 35 40 45
 Thr Glu Leu Val Leu Ser Asp Cys Met Leu Ser Glu Glu Gly Ala Thr
 50 55 60
 Leu Leu Leu Arg Gly Leu Cys Ala Asn Thr Val Leu Arg Phe Leu Asp
 65 70 75 80
 Leu Lys Gly Asn Asn Leu Arg Ala Ala Gly Ala Glu Ala Leu Gly Lys
 85 90 95
 Leu Leu Gln Gln Asn Lys Ser Ile Gln Ser Leu Thr Leu Glu Trp Asn
 100 105 110
 Ser Leu Gly Thr Trp Asp Asp Ala Phe Ala Thr Phe Cys Gly Gly Leu
 115 120 125

<210> 441
 <211> 364
 <212> DNA
 <213> Homo sapiens

<400> 441
 gcccagtact acgtgaacat gttc gatgcc gagcagggct tcttcgacag gcgcagcccc
 60
 ggcggcgagt tccaagccgg cttggatccg gaatcctggg gcggtctggt cactgagacc
 120
 gacggttgga acttcgcctt ccacgctcca caggacggcc gggggctggc cgcgctctac
 180
 ggcggtccga aaggcttgga gaacaagctc gatgcctttt tcgcgacgcc ggaaaacgcg
 240
 gacaagccgg cgtacggcgg aatccacgaa atggtcgagg ccagagcggg ccggatgggc
 300
 caattgggca tgtccaacga gccctcgcac catattcctt acatctacaa ctatgccggc
 360
 gcgc
 364

<210> 442
 <211> 121
 <212> PRT
 <213> Homo sapiens

<400> 442
 Ala Gln Tyr Tyr Val Asn Met Phe Asp Ala Glu Gln Gly Phe Phe Asp
 1 5 10 15
 Arg Arg Ser Pro Gly Gly Glu Phe Gln Ala Gly Leu Asp Pro Glu Ser
 20 25 30
 Trp Gly Gly Leu Phe Thr Glu Thr Asp Gly Trp Asn Phe Ala Phe His


```

      35              40              45
Ala Pro Gln Asp Gly Arg Gly Leu Ala Ala Leu Tyr Gly Gly Pro Lys
   50              55              60
Gly Leu Glu Asn Lys Leu Asp Ala Phe Phe Ala Thr Pro Glu Asn Ala
65              70              75              80
Asp Lys Pro Ala Tyr Gly Gly Ile His Glu Met Val Glu Ala Arg Ala
      85              90              95
Val Arg Met Gly Gln Leu Gly Met Ser Asn Glu Pro Ser His His Ile
      100              105              110
Pro Tyr Ile Tyr Asn Tyr Ala Gly Ala
      115              120

```

<210> 443
 <211> 430
 <212> DNA
 <213> Homo sapiens

```

<400> 443
accgggttacg gctcagtgca acaagagatg ttcgccaaca acctcgtgcg gatgccgctg
60
ctcatgggtgc tggcaatccc cttcgccaag atcctctcga cgacctgtc catcggatcg
120
ggcgggtccgg cggcgtcttc cggccctggc atgggtcatcg gcggagccac tggcgcgga
180
ctgtggcgcc tcctcgaggg gctgccaggt atcccatcct caccgatgag ttctgtcatt
240
gtcggcatga tcgcctgctt cggtgcggtt gcccatgcc cactcggcgt gctgctcatg
300
gttggcgaga tgaccggaaa cctgtcgctg ctcgctcctg gcatgatcgc cgtcgccgtc
360
gctggccgag ttgtcgggga cacttcgatc tacacctctc agctcaagga tcgcctggag
420
ggcgacgcgt
430

```

<210> 444
 <211> 143
 <212> PRT
 <213> Homo sapiens

```

<400> 444
Thr Gly Tyr Gly Ser Val Gln Gln Glu Met Phe Ala Asn Asn Leu Val
 1              5              10              15
Arg Met Pro Leu Leu Met Val Leu Ala Ile Pro Phe Ala Lys Ile Leu
      20              25              30
Ser Thr Thr Leu Ser Ile Gly Ser Gly Gly Pro Ala Ala Ser Ser Gly
      35              40              45
Pro Gly Met Val Ile Gly Gly Ala Thr Gly Ala Ala Leu Trp Arg Leu
      50              55              60
Leu Glu Gly Leu Pro Gly Ile Pro Ser Ser Pro Met Ser Phe Val Ile
65              70              75              80
Val Gly Met Ile Ala Cys Phe Gly Ala Val Ala His Ala Pro Leu Gly
      85              90              95
Val Leu Leu Met Val Gly Glu Met Thr Gly Asn Leu Ser Leu Leu Ala

```

| | | | | | |
|---|-----|--|-----|--|-----|
| | 100 | | 105 | | 110 |
| Pro Gly Met Ile Ala Val Ala Val Ala Gly Arg Val Val Gly Asp Thr | | | | | |
| | 115 | | 120 | | 125 |
| Ser Ile Tyr Thr Ser Gln Leu Lys Asp Arg Leu Glu Gly Asp Ala | | | | | |
| | 130 | | 135 | | 140 |

<210> 445
 <211> 360
 <212> DNA
 <213> Homo sapiens

<400> 445
 ccattggggct gcctagcctc tggggaggcc cctcagctgg tgacaccagc agggcagatt
 60
 tcttgcttta ttgctcacc tgtccagggt tccctctggt tgtgaggag ctgctgccac
 120
 cttgggtcca ggaagcatga agctccgcag gtcagcctcc tgggtggagg acttttcctt
 180
 agttttcttt gctcttctgc tctgagtcga gccctggctg gacctttgat cccttctctc
 240
 tttatcagga aattttctga ctttcttctt ttgccttttc aagatctgtg atgccatctc
 300
 caagtgggaa caagccatga aggagctgca ccccgaaag tctgagggtg ggacacgcgt
 360

<210> 446
 <211> 101
 <212> PRT
 <213> Homo sapiens

| | | | | | |
|---|-----|----|----|--|--|
| Met Ala Cys Ser His Leu Glu Met Ala Ser Gln Ile Leu Lys Arg Gln | | | | | |
| 1 | 5 | 10 | 15 | | |
| Lys Lys Lys Val Arg Lys Phe Pro Asp Lys Glu Arg Arg Asp Gln Arg | | | | | |
| | 20 | 25 | 30 | | |
| Ser Ser Gln Gly Trp Thr Gln Ser Arg Arg Ala Lys Lys Thr Lys Glu | | | | | |
| | 35 | 40 | 45 | | |
| Lys Ser Ser His Gln Glu Ala Asp Leu Arg Ser Phe Met Leu Pro Gly | | | | | |
| 50 | 55 | 60 | | | |
| Pro Lys Val Ala Ala Ala Pro Ser Gln Thr Glu Gly Thr Leu Asp Arg | | | | | |
| 65 | 70 | 75 | 80 | | |
| Val Ser Asn Lys Ala Arg Asn Leu Pro Cys Trp Cys His Gln Leu Arg | | | | | |
| | 85 | 90 | 95 | | |
| Gly Leu Pro Arg Gly | | | | | |
| | 100 | | | | |

<210> 447
 <211> 487
 <212> DNA
 <213> Homo sapiens

<400> 447
 acgcgtgaag ggggaaattg ctcgtgccac ctgaggatta atcattaccc tggaaccctt
 60

cccaaggcca tcaaggaaca cgcacccctt accagacctt ccagctgctg ggggctctcc
 120
 gagtgaggct gaggtcatgg agaaggggaat ggggggcccc catggccagc tggacctgat
 180
 cactgcctcc cactcagcc acagccctca gggccctgtg ccagtccaga agcccattca
 240
 gggacacctt tggccaatgt tctgtttcat ctgcgaggca accttcccca gtgccccaac
 300
 catagcgttt tcccccaaac accctcagga aggagggacc actacctgtg cagggggggc
 360
 caggagcctc ctgagagcct catatgggga ggaagtggta ccatctcacc cccattgcct
 420
 ttctctccta ctccacctg gccagcttcc ctcagtggcc ctectgcctc agtgccccctt
 480
 cacgcgt
 487

<210> 448
 <211> 117
 <212> PRT
 <213> Homo sapiens

<400> 448
 Met Glu Lys Gly Met Gly Gly Pro His Gly Gln Leu Asp Leu Ile Thr
 1 5 10 15
 Ala Ser Pro Leu Ser His Ser Pro Gln Gly Pro Val Pro Val Gln Lys
 20 25 30
 Pro Ile Gln Gly His Leu Trp Pro Met Phe Cys Phe Ile Cys Glu Ala
 35 40 45
 Thr Phe Pro Ser Ala Pro Thr Ile Ala Phe Ser Pro Lys His Pro Gln
 50 55 60
 Glu Gly Gly Thr Thr Thr Cys Ala Gly Gly Ala Arg Ser Leu Leu Arg
 65 70 75 80
 Ala Ser Tyr Gly Glu Val Val Pro Ser His Pro His Cys Leu Ser
 85 90 95
 Leu Leu Leu Pro Pro Gly Gln Leu Pro Ser Val Pro Leu Leu Pro Gln
 100 105 110
 Cys Pro Phe Thr Arg
 115

<210> 449
 <211> 353
 <212> DNA
 <213> Homo sapiens

<400> 449
 gagctcagcc agttggagtt tgagaagcgg cagctgcaca gggacttgga gcaggccaag
 60
 ggaagggggg agcgggcaga gaagctggag agggagctac agcgactcca ggaggagaac
 120
 gggaggctgg ccaggaaggt gacctccctg gagacagcca ccgagaaagt cgaggccctg
 180
 gagcatgaga gccaggcct gcagctggag aaccggactc tgaggaagtc tctggacacc
 240

ttgcagaacg tgccctgca gcttgagggc ctggagcgtg acaacaagca gctggacgca
 300
 gagaacctgg agctgcgcag gctggtggag accatgcgga gacgacaacg cgt
 353

<210> 450
 <211> 117
 <212> PRT
 <213> Homo sapiens

<400> 450
 Glu Leu Ser Gln Leu Glu Phe Glu Lys Arg Gln Leu His Arg Asp Leu
 1 5 10 15
 Glu Gln Ala Lys Glu Lys Gly Glu Arg Ala Glu Lys Leu Glu Arg Glu
 20 25 30
 Leu Gln Arg Leu Gln Glu Glu Asn Gly Arg Leu Ala Arg Lys Val Thr
 35 40 45
 Ser Leu Glu Thr Ala Thr Glu Lys Val Glu Ala Leu Glu His Glu Ser
 50 55 60
 Gln Gly Leu Gln Leu Glu Asn Arg Thr Leu Arg Lys Ser Leu Asp Thr
 65 70 75 80
 Leu Gln Asn Val Ser Leu Gln Leu Glu Gly Leu Glu Arg Asp Asn Lys
 85 90 95
 Gln Leu Asp Ala Glu Asn Leu Glu Leu Arg Arg Leu Val Glu Thr Met
 100 105 110
 Arg Arg Arg Gln Arg
 115

<210> 451
 <211> 444
 <212> DNA
 <213> Homo sapiens

<400> 451
 gtgatgcggc tgactaagcc tactttattc accaatatcc cagtaacatg tgaagagaaa
 60
 gacttacctg gagatctctt taaccagctg atgagagatg atccttcaac cgттаатггт
 120
 gcagaagttt таатгттггг agaaatgctg actttaccac agaattttgg gaatatattt
 180
 ttgggagaga ctttttccag ttatatcagc gttcataatg atagcaatca agttgtaaaa
 240
 gacatattag taaaagctga tcttcagaca agttctcagc gtttaaattct ttcagcctcc
 300
 aatgctgcag tggctgaact taaaccggat tgttgtattg atgatgtcat acatcatgaa
 360
 gtcaaagaaa ttggaacaca catcttggtg tgtgctgtga gttatacaac tcaggctgga
 420
 gaaaaaatgt atttcagaaa attt
 444

<210> 452
 <211> 148
 <212> PRT

<213> Homo sapiens

<400> 452

Val Met Arg Leu Thr Lys Pro Thr Leu Phe Thr Asn Ile Pro Val Thr
 1 5 10 15
 Cys Glu Glu Lys Asp Leu Pro Gly Asp Leu Phe Asn Gln Leu Met Arg
 20 25 30
 Asp Asp Pro Ser Thr Val Asn Gly Ala Glu Val Leu Met Leu Gly Glu
 35 40 45
 Met Leu Thr Leu Pro Gln Asn Phe Gly Asn Ile Phe Leu Gly Glu Thr
 50 55 60
 Phe Ser Ser Tyr Ile Ser Val His Asn Asp Ser Asn Gln Val Val Lys
 65 70 75 80
 Asp Ile Leu Val Lys Ala Asp Leu Gln Thr Ser Ser Gln Arg Leu Asn
 85 90 95
 Leu Ser Ala Ser Asn Ala Ala Val Ala Glu Leu Lys Pro Asp Cys Cys
 100 105 110
 Ile Asp Asp Val Ile His His Glu Val Lys Glu Ile Gly Thr His Ile
 115 120 125
 Leu Val Cys Ala Val Ser Tyr Thr Thr Gln Ala Gly Glu Lys Met Tyr
 130 135 140
 Phe Arg Lys Phe
 145

<210> 453

<211> 373

<212> DNA

<213> Homo sapiens

<400> 453

gctagctctg accccacctt tgccaagtgg cactaggggtg gccaatgggg actaggggtg
 60
 tataattgga aaatacagtc tccctgttg tccaagaaag gcccagatg acctggggct
 120
 tgaaaggcac tcccgctggg tgcttcctgg gagcaggtgg ggggcagcgg ggcggcgggg
 180
 cctgtctgtg ctgagcatcc ccagctccag ggcaggtgct gggctctgag cccactggt
 240
 gcgttttggg atgggctggc ctgcgcggct gtcgtttcag agcacacaga agagaccctg
 300
 ccacaggagg agtgggagga gaagctgttg atgttcctgc gagacaccct ggccatcatt
 360
 tctgacaacg cgt
 373

<210> 454

<211> 108

<212> PRT

<213> Homo sapiens

<400> 454

Met Met Ala Arg Val Ser Arg Arg Asn Ile Asn Ser Phe Ser Ser His
 1 5 10 15
 Ser Ser Cys Gly Arg Val Ser Ser Val Cys Ser Glu Thr Thr Ala Ala

20 25 30
 Gln Ala Ser Pro Ser Gln Asn Ala Pro Val Gly Leu Arg Ala Gln His
 35 40 45
 Leu Pro Trp Ser Trp Gly Cys Ser Ala Gln Thr Gly Pro Ala Ala Pro
 50 55 60
 Leu Pro Pro Thr Cys Ser Gln Glu Ala Pro Ser Gly Ser Ala Phe Gln
 65 70 75 80
 Ala Pro Gly His Leu Gly Pro Phe Leu Asp Asn Arg Gly Asp Cys Ile
 85 90 95
 Phe Gln Leu Tyr Asn Pro Ser Pro His Trp Pro Pro
 100 105

<210> 455

<211> 602

<212> DNA

<213> Homo sapiens

<400> 455

cctaggcaaa gcatgcccac cctacctccc cttaccctta cccttcattt tcccctaagc
 60
 acccatcacc accgatgtta ctgtatgtgt ttgcttacgc tgacagccca ccaccacac
 120
 tggaaatgcc gcacgacaaa ggcaggactc ttggctgcct tagccacagc tggatcccca
 180
 gagcttttga ggggtgttggg cacagagtgg agtgggtact taataagtat ctgtggaatg
 240
 aacatgtaca gagtgaagcc ctgtgcccag aacaggetca aaataagetc aattcctttc
 300
 cttgccactt actaagtcct ttttctctcg cccctctca ctgacctggg tttgatgccca
 360
 gacagcacag atgggctagg gaggcagggtg gggaagcaga gatctgcgtc tcttgagct
 420
 ggagctgggt ggtggggctc cttcctgggt ctgcgagggc tcattgggga ggtggcagcg
 480
 acccctcag gagcctctgt cgctgcact cagatctgtg cctttccaca gcgcccggag
 540
 gaagacttgc tcaggagata aattcaaaga caacaggaag ctggacgtgg tggctcacgc
 600
 gt
 602

<210> 456

<211> 100

<212> PRT

<213> Homo sapiens

<400> 456

Met Pro Thr Leu Pro Pro Leu Thr Leu Thr Leu His Phe Pro Leu Ser
 1 5 10 15
 Thr His His His Arg Cys Tyr Cys Met Cys Leu Leu Thr Leu Thr Ala
 20 25 30
 His His Pro His Trp Asn Val Arg Thr Thr Lys Ala Gly Leu Leu Ala
 35 40 45
 Ala Leu Ala Thr Ala Gly Ser Pro Glu Leu Cys Arg Val Leu Gly Thr

```

      50              55              60
Glu Trp Ser Gly Tyr Leu Ile Ser Ile Cys Gly Met Asn Met Tyr Arg
65              70              75              80
Val Lys Pro Cys Ala Gln Asn Arg Leu Lys Ile Ser Ser Ile Pro Phe
      85              90              95
Leu Ala Thr Tyr
      100

```

<210> 457
 <211> 324
 <212> DNA
 <213> Homo sapiens

```

<400> 457
acgcgtcatg tggatattcc tgggaggttc ccaggaacgt ttctggacgg gccccgacc
60
agaggtcagg gaacttttct tattattctg cactgcca gggatagtca aaccaggtct
120
tccccctctg ctggccgcaa cagccagcc gccgccacga ccgcacgctg aattcatgac
180
ccgacacgcg acgtggcagc gagcacaccc accgctagga gaaagagcgc tcatcgaaga
240
tcgttttctg tccactggcc agcgccacta tgatcagggtg gggatatccgc ccggcgggcg
300
gagcaccggg acgccggggc gccg
324

```

<210> 458
 <211> 105
 <212> PRT
 <213> Homo sapiens

```

<400> 458
Met Trp Ile Phe Leu Gly Gly Ser Gln Glu Arg Phe Trp Thr Gly Pro
1      5      10      15
Arg Pro Glu Val Arg Glu Leu Phe Leu Leu Phe Cys Thr Cys Pro Gly
      20      25      30
Ile Val Lys Pro Gly Leu Pro Leu Leu Leu Ala Ala Thr Arg Gln Pro
      35      40      45
Pro Pro Arg Pro His Ala Glu Phe Met Thr Arg His Ala Thr Trp Gln
      50      55      60
Arg Ala His Pro Pro Leu Gly Glu Arg Ala Leu Ile Glu Asp Arg Phe
65      70      75      80
Leu Ser Thr Gly Gln Arg His Tyr Asp Gln Val Gly Tyr Pro Pro Gly
      85      90      95
Gly Gly Ser Thr Gly Thr Pro Gly Arg
      100      105

```

<210> 459
 <211> 415
 <212> DNA
 <213> Homo sapiens

<400> 459

acgcgttcat tggcatctg cttccatgga tttcctgcgg ggaggcgcg cgcagagtgc
 60
 ggggtgctgaa cagcagactt cagtgatcgt ttcaaccacc ggccgagatg ggtcctgacg
 120
 ctgggcttca agccgcttgc gctcgcgctc ctgatctcgg gcagcgcgat tccgggtggtt
 180
 tatgctgccg gcagacgact gcgcacgccc ctcacgaggt atctgcacat gcttaaaggg
 240
 agaggcctca cccgacagct gggcatcgga tttacgaagc ccacgacgaa tttcctcgc
 300
 ctctcaaag ccgatcatcg gcatgccagg tttgtggttg aatgcttcga tcaacacact
 360
 aggatcggtg ggggtccacca catacaccga gcggcaatcg agcggatacg acctc
 415

<210> 460

<211> 105

<212> PRT

<213> Homo sapiens

<400> 460

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Pro | Met | Ile | Gly | Phe | Glu | Glu | Ala | Arg | Lys | Ile | Arg | Arg | Gly | Leu |
| 1 | | | | 5 | | | | | 10 | | | | | 15 | |
| Arg | Lys | Ser | Asp | Ala | Gln | Leu | Ser | Gly | Glu | Ala | Ser | Pro | Phe | Lys | His |
| | | 20 | | | | | | 25 | | | | 30 | | | |
| Val | Gln | Ile | Pro | Arg | Glu | Gly | Arg | Ala | Gln | Ser | Ser | Ala | Gly | Ser | Ile |
| | | 35 | | | | | 40 | | | | | 45 | | | |
| Asn | His | Arg | Asn | Arg | Ala | Ala | Arg | Asp | Gln | Glu | Arg | Glu | Arg | Lys | Arg |
| | 50 | | | | | 55 | | | | 60 | | | | | |
| Leu | Glu | Ala | Gln | Arg | Gln | Asp | Pro | Ser | Arg | Pro | Val | Val | Glu | Thr | Ile |
| 65 | | | | 70 | | | | | 75 | | | | | 80 | |
| Thr | Glu | Val | Ser | Cys | Ser | Thr | Pro | Ala | Leu | Ser | Ala | Ala | Pro | Pro | Arg |
| | | | 85 | | | | | | 90 | | | | | 95 | |
| Arg | Lys | Ser | Met | Glu | Ala | Asp | Ala | Glu | | | | | | | |
| | | | 100 | | | | | 105 | | | | | | | |

<210> 461

<211> 357

<212> DNA

<213> Homo sapiens

<400> 461

acgcgttcga ggtcggctaa atttatcatg cgcacgacaa agagagtagt ggctcacaac
 60
 cgggtcacat gcatgatgac aaaaactggc agaataagagt tgatgtcatc ccgtctacca
 120
 gctcctagaa ccagctcaga gagtcccggt gtcggtaccg tcgagactca gtacacaact
 180
 gtcgcgatac cggacgaccc tttcatctg gttgcagatg ggcgtctcaa tcacgtcact
 240
 gtcgcttacg aaacctacgg gaagctcaat acgtccagcg acaatgcggt ctatacctgt
 300
 catgcgctta ctggtgatgc ccatgcagcc ggatttcacc ccggtgtagt ccgtccg
 357

<210> 462
 <211> 119
 <212> PRT
 <213> Homo sapiens

<400> 462
 Thr Arg Ser Arg Ser Ala Lys Phe Ile Met Arg Thr Thr Lys Arg Val
 1 5 10 15
 Val Ala His Asn Arg Val Thr Cys Met Met Thr Lys Thr Gly Arg Ile
 20 25 30
 Glu Leu Met Ser Ser Arg Leu Pro Ala Pro Arg Thr Ser Ser Glu Ser
 35 40 45
 Pro Gly Val Gly Thr Val Glu Thr Gln Tyr Thr Thr Val Ala Ile Pro
 50 55 60
 Asp Asp Pro Leu His Leu Val Ala Asp Gly Arg Leu Asn His Val Thr
 65 70 75 80
 Val Ala Tyr Glu Thr Tyr Gly Lys Leu Asn Thr Ser Ser Asp Asn Ala
 85 90 95
 Val Tyr Thr Cys His Ala Leu Thr Gly Asp Ala His Ala Ala Gly Phe
 100 105 110
 His Pro Gly Val Val Arg Pro
 115

<210> 463
 <211> 434
 <212> DNA
 <213> Homo sapiens

<400> 463
 gtgcacgggg tatgcgaggg atgcggcatt gccaccaatg ccgctgacct gcgcagatac
 60
 gaggcagctg gtgacgatga agtggtgcga tgcgaggaat gcgatcgtat cctggtgcgt
 120
 accggagagt ccatctgagc ccttcttggtg gcggtgatgc cgggatatcc gtagaattag
 180
 cggtcggacg agccatccgg gtgatcgagg cagcgggtgag ttgtcgagga aagtcggggc
 240
 tccatagagc aggggtggtgg gtaacgcccc cccgggggtga cccgcgggaa agtgccacag
 300
 agaacagact gccggtttcg agccgggtgag ggtgaaacgg tggagtaagt gccaccgcg
 360
 tcatcggtga cggtgacggc atggcaaacc ccacctggag caaggccaag aagaccgtga
 420
 ggtcgcggaac gcgt
 434

<210> 464
 <211> 127
 <212> PRT
 <213> Homo sapiens

<400> 464
 Met Pro Ser Pro Ser Pro Met Thr Arg Trp Ala Leu Thr Pro Pro Phe

```

      1           5           10           15
His Pro His Arg Leu Glu Thr Gly Ser Leu Phe Ser Val Ala Leu Ser
      20           25           30
Arg Gly Ser Pro Arg Val Gly Val Thr His His Pro Ala Leu Trp Ser
      35           40           45
Pro Asp Phe Pro Arg Gln Leu Thr Ala Ala Ala Ile Thr Arg Met Ala
      50           55           60
Arg Pro Thr Ala Asn Ser Thr Asp Ile Pro Ala Ser Pro Pro Gln Glu
      65           70           75           80
Gly Leu Arg Trp Thr Leu Arg Tyr Ala Pro Gly Tyr Asp Arg Ile Pro
      85           90           95
Arg Ile Ala Pro Leu His Arg His Gln Leu Pro Arg Ile Cys Ala Gly
      100          105          110
Gln Arg His Trp Trp Gln Cys Arg Ile Pro Arg Ile Pro Arg Ala
      115          120          125

```

<210> 465
 <211> 438
 <212> DNA
 <213> Homo sapiens

```

<400> 465
gatcatttag aatttatgga agaagctgat gtgaaagcta tggtaaatac tggcactgtg
60
gctgtattgc taccaggagc attttacacc ttgaaagaaa ctcaacttcc accgatgaat
120
ttgttacgtc agtacggagt agacattgct atttcgacgg atgctaatac agggacgtcg
180
ccagcgttat cattacggtt aatgatgaat atggcatgta ccttgtttgg tatgacacct
240
gaaaccgccc ttgcaggggt aacaattcat gcggcaaaag cgttggggat tagcgattct
300
catggcactt tagaagttgg caaggtagct gattttgtct gctgggatgt ggaaagcccc
360
ggtgaacttt gttattggtt aggagagcag ttagtaaagc aacgtattca gcacggagta
420
tcccatgaat aatctaga
438

```

<210> 466
 <211> 143
 <212> PRT
 <213> Homo sapiens

```

<400> 466
Asp His Leu Glu Phe Met Glu Glu Ala Asp Val Lys Ala Met Val Lys
      1           5           10           15
Ser Gly Thr Val Ala Val Leu Leu Pro Gly Ala Phe Tyr Thr Leu Lys
      20           25           30
Glu Thr Gln Leu Pro Pro Met Asn Leu Leu Arg Gln Tyr Gly Val Asp
      35           40           45
Ile Ala Ile Ser Thr Asp Ala Asn Pro Gly Thr Ser Pro Ala Leu Ser
      50           55           60
Leu Arg Leu Met Met Asn Met Ala Cys Thr Leu Phe Gly Met Thr Pro

```

```

65          70          75          80
Glu Thr Ala Leu Ala Gly Val Thr Ile His Ala Ala Lys Ala Leu Gly
          85          90          95
Ile Ser Asp Ser His Gly Thr Leu Glu Val Gly Lys Val Ala Asp Phe
          100          105          110
Val Cys Trp Asp Val Glu Ser Pro Gly Glu Leu Cys Tyr Trp Leu Gly
          115          120          125
Glu Gln Leu Val Lys Gln Arg Ile Gln His Gly Val Ser His Glu
          130          135          140

```

<210> 467
 <211> 460
 <212> DNA
 <213> Homo sapiens

```

<400> 467
ntttccctgg ctattggcca tgtgggacac aacgttcgcg ctaccccaga gcggttaagc
60
tgcacccctg caccttcttc tcccaccgct tcaaagccac agtgaggaac ttcggagctt
120
ctcgcagtga agatggcggt ggaggaatgg atgccctggc tagaagaggc ggaatatctg
180
ttgattgtgt ggaccgacca caaaaacctg gagtatctcc acacaaccaa gtgcctcaac
240
tccaggcaag caagaagggc ccagctgttt acctggttcc acttttccct ctccctaccg
300
ccgggggtcca agaacatcag gctggatgcc ctttcttgcc actttatggg catggggcca
360
ttctccagg cttgcctgtc acccgggctc ccgtcaaacc ctggccttcg tgcgacaaca
420
ctcttggtgc cttctatggt tctgtatggt gccgcaattg
460

```

<210> 468
 <211> 118
 <212> PRT
 <213> Homo sapiens

```

<400> 468
Gly Thr Ser Glu Leu Leu Ala Val Lys Met Ala Leu Glu Glu Trp Met
1      5      10      15
Pro Trp Leu Glu Glu Ala Glu Tyr Leu Leu Ile Val Trp Thr Asp His
20     25     30
Lys Asn Leu Glu Tyr Leu His Thr Thr Lys Cys Leu Asn Ser Arg Gln
35     40     45
Ala Arg Arg Ala Gln Leu Phe Thr Trp Phe His Phe Ser Leu Ser Tyr
50     55     60
Arg Pro Gly Ser Lys Asn Ile Arg Leu Asp Ala Leu Ser Cys His Phe
65     70     75     80
Met Gly Met Gly Pro Phe Leu Gln Ala Cys Leu Ser Pro Gly Leu Pro
85     90     95
Ser Asn Pro Gly Leu Arg Ala Thr Thr Leu Leu Val Pro Ser Met Val
100    105    110
Leu Tyr Val Ala Ala Ile

```

115

<210> 469
 <211> 381
 <212> DNA
 <213> Homo sapiens

<400> 469
 cttgtgcaca cggtattttt ccaatacaaa tagtttaaaa agtaaactcc aaatacctat
 60
 aagccccctc aaagcacctt ccaaatatga accttggtta tgcccaagggt ccagaggggt
 120
 cccccagaaa ggcccaggag cctggggcat gggaaagctg tcgggggtccc catgctgact
 180
 ccctggactc caagcgatat tccataaagc cagggcctcc tggctgcggg agggaggcct
 240
 tgacccaaaa tccattcggc cctggatact ggagaggcag aggcctctgc tgatgagaag
 300
 ccctgagttc ctggctagct gtgggtaacc acaaaaaatg cgggggggtga tgattttcga
 360
 agtccatcgg caaagaaaga c
 381

<210> 470
 <211> 110
 <212> PRT
 <213> Homo sapiens

<400> 470
 Met Asp Phe Glu Asn His His Pro Pro His Phe Leu Trp Leu Thr Thr
 1 5 10 15
 Ala Ser Gln Glu Leu Arg Ala Ser His Gln Gln Arg Pro Leu Pro Leu
 20 25 30
 Gln Tyr Pro Gly Pro Asn Gly Phe Trp Val Lys Ala Ser Leu Pro Gln
 35 40 45
 Pro Gly Gly Pro Gly Phe Met Glu Tyr Arg Leu Glu Ser Arg Glu Ser
 50 55 60
 Ala Trp Gly Pro Arg Gln Leu Ser His Ala Pro Gly Ser Trp Ala Phe
 65 70 75 80
 Leu Gly Asp Pro Ser Gly Pro Trp Ala Leu Thr Arg Phe Ile Phe Gly
 85 90 95
 Arg Cys Phe Glu Gly Ala Tyr Arg Tyr Leu Glu Phe Thr Phe
 100 105 110

<210> 471
 <211> 378
 <212> DNA
 <213> Homo sapiens

<400> 471
 accggtgact acctgcagca ctggattgac atgggtaaaa agggcggcga ccgcatgcc
 60
 gaggtcttcc tgggtaactg gttccgccgc ggcgacgatg gccgcttcct gtggccgngg
 120

cttggcgaaa acttccccgt cctanagtgg atcatcgacc gcattgaagg caacgtagag
 180
 gccgaggaca cggtaggtcgg acgcaccgcc cgcgccgagg acatcgactt gcaaggcctt
 240
 gacttcgatg tcgacgacgt tcgcgcgcga ctgcgcgttg acccgaagga atgggaaggc
 300
 gatatgcaag acaacgccga gtacctgaac ttcttgggct cccgcgtgcc cgaggaagtg
 360
 tggaaccagt tccgcgcc
 378

<210> 472

<211> 126

<212> PRT

<213> Homo sapiens

<400> 472

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Thr | Gly | Asp | Tyr | Leu | Gln | His | Trp | Ile | Asp | Met | Gly | Lys | Lys | Gly | Gly |
| 1 | | | | 5 | | | | 10 | | | | | | 15 | |
| Asp | Arg | Met | Pro | Glu | Val | Phe | Leu | Val | Asn | Trp | Phe | Arg | Arg | Gly | Asp |
| | | 20 | | | | | 25 | | | | | 30 | | | |
| Asp | Gly | Arg | Phe | Leu | Trp | Pro | Xaa | Leu | Gly | Glu | Asn | Phe | Pro | Val | Leu |
| | 35 | | | | | 40 | | | | | 45 | | | | |
| Xaa | Trp | Ile | Ile | Asp | Arg | Ile | Glu | Gly | Asn | Val | Glu | Ala | Glu | Asp | Thr |
| 50 | | | | | 55 | | | | | 60 | | | | | |
| Val | Val | Gly | Arg | Thr | Ala | Arg | Ala | Glu | Asp | Ile | Asp | Leu | Gln | Gly | Leu |
| 65 | | | 70 | | | | | 75 | | | | | 80 | | |
| Asp | Phe | Asp | Val | Asp | Asp | Val | Arg | Ala | Ala | Leu | Ala | Val | Asp | Pro | Lys |
| | | 85 | | | | | | 90 | | | | | 95 | | |
| Glu | Trp | Glu | Gly | Asp | Met | Gln | Asp | Asn | Ala | Glu | Tyr | Leu | Asn | Phe | Leu |
| | | 100 | | | | | 105 | | | | | 110 | | | |
| Gly | Ser | Arg | Val | Pro | Glu | Glu | Val | Trp | Asn | Gln | Phe | Arg | Ala | | |
| | | 115 | | | | | 120 | | | | | 125 | | | |

<210> 473

<211> 339

<212> DNA

<213> Homo sapiens

<400> 473

accggttggt gggggaaggg acccatccca tgccacctgt cctagaaaat gtttcccctt
 60
 gttgagcagc tgctggatct agggctgctg ggtctaagtc caaaaaggga aaaaggaaaa
 120
 aggcaccaag taaaagaagg gggaagctgc caaaaccccc cctgcaaaa ctctcccacc
 180
 ctgcttccat ttcctctcc agggacaggtgtgtacctccc ctctccctg tctctctcag
 240
 atgccccagg ggtctcttac ttcattctctg ccgacctgc caggagtggc ctcaggggta
 300
 gaggctccta gttggagaat ttgcttgacg gaaggtgaa
 339

<210> 474

<211> 97
 <212> PRT
 <213> Homo sapiens

<400> 474
 Met Phe Pro Leu Val Glu Gln Leu Leu Asp Leu Gly Leu Leu Gly Leu
 1 5 10 15
 Ser Pro Lys Arg Glu Lys Gly Lys Arg His Gln Val Lys Glu Gly Gly
 20 25 30
 Ser Cys Gln Asn Pro Pro Cys Gln Asn Ser Pro Thr Leu Leu Pro Phe
 35 40 45
 Pro Ser Pro Gly Asn Arg Cys Thr Ser Pro Pro Pro Cys Pro Pro Gln
 50 55 60
 Met Pro Gln Gly Leu Ser Thr Ser Phe Leu Pro Thr Leu Pro Gly Val
 65 70 75 80
 Ala Ser Gly Val Glu Ala Pro Ser Trp Arg Ile Cys Leu Gln Glu Gly
 85 90 95
 Glu

<210> 475
 <211> 345
 <212> DNA
 <213> Homo sapiens

<400> 475
 acgcgtgaag ggtcccoctcc aaactctgag cctccttcca agccttgctg ggagctcccc
 60
 agcgctgcc ggagaggcct ctctccagg cgggcttccc gcgccgatgt gaaggagagg
 120
 ctgccccaga ggggtctgga tcgtaatcca gaaagggaca gtccacacagc cataatcccc
 180
 aatgctggga ctcttcagta aaggaagaga tggttttttc gttcatctgc ctttctgaaa
 240
 ggtaaaatat ctccagatcc gggctctctg ggcgactgcg tatgtggggg tccctgaagc
 300
 ctttgatgga tcttggttaga agtgggttgt tcatcttggg gtttt
 345

<210> 476
 <211> 111
 <212> PRT
 <213> Homo sapiens

<400> 476
 Met Asn Asn Pro Leu Leu Thr Arg Ser Ile Lys Gly Phe Arg Asp Pro
 1 5 10 15
 His Ile Arg Ser Arg Pro Glu Ser Pro Asp Leu Glu Ile Phe Tyr Leu
 20 25 30
 Ser Glu Arg Gln Met Asn Glu Lys Ala Ile Ser Ser Phe Thr Glu Glu
 35 40 45
 Ser Gln His Ser Gly Leu Trp Leu Trp Asp Cys Pro Phe Leu Asp Tyr
 50 55 60
 Asp Pro Asp Pro Ser Gly Ala Ala Ser Pro Ser His Arg Arg Gly Lys

```

65          70          75          80
Pro Ala Trp Arg Arg Gly Leu Ser Gly Arg Arg Trp Gly Ala Pro Ser
          85          90          95
Lys Ala Trp Lys Glu Ala Gln Ser Leu Glu Gly Thr Leu His Ala
          100          105          110

```

<210> 477
 <211> 422
 <212> DNA
 <213> Homo sapiens

```

<400> 477
acgcgtggcc gagccagcgt gctcaaggaa atggtcaacg gcactcttat taacggctgg
60
gactctcccg aggtggaacg ggcactggac ctgtgcatgg cgtgcaaagg gtgcgcccga
120
gattgccccca ccggaatcga catggccagc taccgcagca cggttcttga cgaaaaatac
180
cgtcaccgtc tccgccctcg ctcccacctg acgatggggc tgctgcccac gtgggaacgt
240
ttgctcaatc ggaccccagg agcgccgtcg ctggctaacg cagtgttttc gatgccggtc
300
ttcgcacgtc ttgctagatg gacagccggg gtggatcagc gtcgtcccct cccccgattc
360
cagccctcgg ccagattggc cagtccgcag gccgccccgg ttaaggagat tgtggcggat
420
cc
422

```

<210> 478
 <211> 140
 <212> PRT
 <213> Homo sapiens

```

<400> 478
Thr Arg Gly Arg Ala Ser Val Leu Lys Glu Met Val Asn Gly Thr Leu
1          5          10          15
Ile Asn Gly Trp Asp Ser Pro Glu Val Glu Arg Ala Leu Asp Leu Cys
          20          25          30
Met Ala Cys Lys Gly Cys Ala Arg Asp Cys Pro Thr Gly Ile Asp Met
          35          40          45
Ala Ser Tyr Arg Ser Thr Val Leu Asp Glu Lys Tyr Arg His Arg Leu
          50          55          60
Arg Pro Arg Ser His Leu Thr Met Gly Leu Leu Pro Met Trp Glu Arg
65          70          75          80
Leu Leu Asn Arg Thr Pro Gly Ala Pro Ser Leu Ala Asn Ala Val Leu
          85          90          95
Ser Met Pro Val Phe Ala Arg Leu Ala Arg Trp Thr Ala Gly Val Asp
          100          105          110
Gln Arg Arg Pro Leu Pro Arg Phe Gln Pro Ser Ala Arg Leu Ala Ser
          115          120          125
Pro Gln Ala Ala Pro Val Lys Glu Ile Val Ala Asp
          130          135          140

```

<210> 479
 <211> 348
 <212> DNA
 <213> Homo sapiens

<400> 479
 cgctgggcca ttggccgggc gctgggtgcgg caccgcgcac tggtgattgc cgatgagccg
 60
 atctcggcgt tggacatgac catccagaag cagattcttg agctgttcga gcgcctgcag
 120
 gcgcagtacg gctttgcctg cctgttcacg tcccacgacc tggcagcggg ggaacgcacg
 180
 gcccaccggg tggcgggtgat gagcgagggc aggggtgggtg aaatgggtgc ccgcgacgag
 240
 atcttcgacc gcccgcagca cccctacacc cgcaagctgc tggccgcgcg cagccccctg
 300
 gagaaacttg aaaacgggtg ctaccgcacg cgccagggcc ccgtaccg
 348

<210> 480
 <211> 116
 <212> PRT
 <213> Homo sapiens

<400> 480
 Arg Val Ala Ile Gly Arg Ala Leu Val Arg His Pro Arg Leu Val Ile
 1 5 10 15
 Ala Asp Glu Pro Ile Ser Ala Leu Asp Met Thr Ile Gln Lys Gln Ile
 20 25 30
 Leu Glu Leu Phe Glu Arg Leu Gln Ala Gln Tyr Gly Phe Ala Cys Leu
 35 40 45
 Phe Ile Ser His Asp Leu Ala Ala Val Glu Arg Ile Ala His Arg Val
 50 55 60
 Ala Val Met Ser Glu Gly Arg Val Val Glu Met Gly Ala Arg Asp Glu
 65 70 75 80
 Ile Phe Asp Arg Pro Gln His Pro Tyr Thr Arg Lys Leu Leu Ala Ala
 85 90 95
 Ala Ser Pro Leu Glu Lys Leu Glu Asn Gly Gly Tyr Arg Ile Arg Gln
 100 105 110
 Gly Pro Val Pro
 115

<210> 481
 <211> 441
 <212> DNA
 <213> Homo sapiens

<400> 481
 aagcttctga ctgtggcatt ctccctgctt aatatgtcct caatatcccc tacttactgg
 60
 gcaaaaatcct gcttatgctt tgggactagc tcaaagacca ctcccttgga tgggtgccttc
 120
 cctgccctgc cggcttgccg tggcttctct agtgtttaga ttaccatcac attgcatcat
 180

gagagcagaa gaccatctcc atgtgactgc tgcccctgct cccagcaggg cccacaanca
 240
 cccagtccag gacctggctc acgctgggtg gcggatgccc aggaatgggg ctctggatct
 300
 gcctcttctc ctgcaggacc aggaaaccgc tgcccctgtcc ctgccccagg aaaccctcag
 360
 taaatcccca gtcatttgag tttcccctca gcgccagaga ccaataacac atctccacca
 420
 acctgaaaaa ccttcacgcg t
 441

<210> 482

<211> 120

<212> PRT

<213> Homo sapiens

<400> 482

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Lys | Leu | Leu | Thr | Val | Ala | Phe | Ser | Leu | Leu | Asn | Met | Ser | Ser | Ile | Ser |
| 1 | | | | 5 | | | | | 10 | | | | | 15 | |
| Pro | Thr | Tyr | Trp | Ala | Lys | Ser | Cys | Leu | Cys | Phe | Gly | Thr | Ser | Ser | Lys |
| | | | 20 | | | | | 25 | | | | | 30 | | |
| Thr | Thr | Pro | Leu | Asp | Gly | Ala | Phe | Pro | Ala | Leu | Pro | Ala | Cys | Ala | Gly |
| | | 35 | | | | 40 | | | | | | 45 | | | |
| Phe | Leu | Ser | Val | Arg | Ile | Thr | Ile | Thr | Leu | His | His | Glu | Ser | Arg | Arg |
| | 50 | | | | 55 | | | | 60 | | | | | | |
| Pro | Ser | Pro | Cys | Asp | Cys | Cys | Pro | Cys | Ser | Gln | Gln | Gly | Pro | Gln | Xaa |
| 65 | | | | 70 | | | | | 75 | | | | | 80 | |
| Pro | Ser | Pro | Gly | Pro | Gly | Ser | Arg | Trp | Val | Ala | Asp | Ala | Gln | Glu | Trp |
| | | | 85 | | | | 90 | | | | | | 95 | | |
| Gly | Ser | Gly | Ser | Ala | Ser | Ser | Pro | Ala | Gly | Pro | Gly | Asn | Arg | Cys | Pro |
| | | | 100 | | | | 105 | | | | | | 110 | | |
| Val | Pro | Ala | Pro | Gly | Asn | Pro | Gln | | | | | | | | |
| | | 115 | | | | 120 | | | | | | | | | |

<210> 483

<211> 330

<212> DNA

<213> Homo sapiens

<400> 483

acgcgttcac tccctgatgg ccacgcacga gctaacggag ggatggggcg aaggaaggc
 60
 caagggttgc tcgaagacca aggagtgtgc agggcaggac ctgcttttaa aggaatatcc
 120
 tctcaccaga gacacgcggc ggccaggcag ggccggagcg gggcctgtgc ccaggctccg
 180
 agcgtctgcc cagcccagca tccctgtccc cagccaggaa tatgtcttcg tggcatagag
 240
 ggagctcttg gagccacacc tgcgtgtgca catgtgtcac cccactgctg ggaggggctc
 300
 tcccgggacc ctgcagcgtg ggctggggcc
 330

<210> 484

<211> 96
 <212> PRT
 <213> Homo sapiens

<400> 484
 Met Gly Arg Arg Glu Gly Gln Gly Cys Leu Glu Asp Gln Gly Val Cys
 1 5 10 15
 Arg Ala Gly Pro Arg Phe Lys Gly Ile Ser Ser His Gln Arg His Ala
 20 25 30
 Ala Ala Arg Gln Gly Arg Ser Gly Ala Cys Ala Gln Ala Pro Ser Val
 35 40 45
 Cys Pro Ala Gln His Pro Cys Pro Gln Pro Gly Ile Cys Leu Arg Gly
 50 55 60
 Ile Glu Gly Ala Leu Gly Ala Thr Pro Ala Cys Ala His Val Ser Pro
 65 70 75 80
 His Cys Trp Glu Gly Leu Ser Arg Asp Pro Ala Ala Trp Ala Gly Pro
 85 90 95

<210> 485
 <211> 377
 <212> DNA
 <213> Homo sapiens

<400> 485
 acgcgtgctc ggcgcggacga agtcggcgct gatcgcccag tcatgcgccc tgcccgtagcc
 60
 gccaggttcg gcgatcgccg cattcggccg gccggaatcg agaaggaatg cgtggacgta
 120
 cgggggatac caaaggaatc ttgtcgaggg cttcgcgggc ctcgacgtgg atcacctgta
 180
 cccgacggac gtggggaagc cgtcccgcga gctcacggga ctccgcgaca tcgatgtgcg
 240
 atacgatttg caccgtcgtc ggctgcgtgc gcgacacatg ctccgcgatc gcctcagcgg
 300
 tggtttccga cgtcagcagg aacgtggcga cgggtggcat ggcggtcgcc gttatgtcgg
 360
 cattcccatt cctcggg
 377

<210> 486
 <211> 111
 <212> PRT
 <213> Homo sapiens

<400> 486
 Met Arg Pro Ala Arg Ala Ala Gln Phe Gly Asp Arg Arg Ile Arg Pro
 1 5 10 15
 Ala Gly Ile Glu Lys Glu Cys Val Asp Val Arg Gly Ile Pro Lys Glu
 20 25 30
 Ser Cys Arg Gly Leu Arg Gly Pro Arg Arg Gly Ser Pro Val Pro Asp
 35 40 45
 Gly Arg Gly Glu Ala Val Pro Gln Ala His Gly Thr Pro Arg His Arg
 50 55 60
 Cys Ala Ile Arg Phe Ala Pro Ser Ser Ala Ala Cys Ala Thr His Ala

```

65          70          75          80
Pro Arg Ser Pro Gln Arg Trp Phe Pro Thr Ser Ala Gly Thr Trp Arg
          85          90          95
Arg Val Ala Trp Arg Ser Pro Leu Cys Arg His Ser His Ser Ser
          100          105          110

```

<210> 487
 <211> 459
 <212> DNA
 <213> Homo sapiens

```

<400> 487
nnacgcgtaa gatcgattgt ggatcagcac cgatgctggt ccccccgcac ttgttgttgg
60
cgggtgttgt tgtaaggagt gtgtgtgatg cgtgttggtg ttcctactga ggtaagaat
120
agtgaatttc gtgtggctgt gacgccggcg ggtgttcattg cgttgggttg tctgggtcat
180
gaggtgttgg ttcaggctgg tgctgggtgtg gggtcgggta ttccggattc ggattttgtg
240
ggtgctggtg cgcgggttgt gggatgatgt gagtcggtgt ggggtgatgc tgatttggtg
300
ttgaaggtga aggagcctgt tgcggaggag tatgggcggt tgcattgaggg tttggttctt
360
tttacgtatc ttcatttggc tgctgatgag gcgttgactc gtgagctttt ggggcgtggg
420
gtgacgtcga ttgcgtatga gacggtggag ttggccgat
459

```

<210> 488
 <211> 124
 <212> PRT
 <213> Homo sapiens

```

<400> 488
Met Arg Val Gly Val Pro Thr Glu Val Lys Asn Ser Glu Phe Arg Val
1          5          10          15
Ala Val Thr Pro Ala Gly Val His Ala Leu Val Gly Arg Gly His Glu
          20          25          30
Val Leu Val Gln Ala Gly Ala Gly Val Gly Ser Gly Ile Pro Asp Ser
          35          40          45
Asp Phe Val Gly Ala Gly Ala Arg Val Val Gly Asp Val Glu Ser Val
          50          55          60
Trp Gly Asp Ala Asp Leu Val Leu Lys Val Lys Glu Pro Val Ala Glu
65          70          75          80
Glu Tyr Gly Arg Leu His Glu Gly Leu Val Leu Phe Thr Tyr Leu His
          85          90          95
Leu Ala Ala Asp Glu Ala Leu Thr Arg Glu Leu Leu Gly Arg Gly Val
          100          105          110
Thr Ser Ile Ala Tyr Glu Thr Val Glu Leu Ala Asp
          115          120

```

<210> 489
 <211> 542

<212> DNA

<213> Homo sapiens

<400> 489

nacgcgtttg gcgtactgag tgcggtggtg gatggcgacg acagtggcaa gccgctgctc
 60
 aaccagcacg gttgctacaa agtgcgcttt ccatttacct gcgatcaaaa gcccagcact
 120
 cgggggttcgg catggctgcg caggggtgctg ttgtctgccg gttccagcca tggcatgcac
 180
 tttccgctgc tcaaaggcag tgaagtgttg gtgtcatttc tggggggcga ccccgaccgg
 240
 ccgattatcg ttggctgcgt accaaactcg gaaaccccgga gcatggctcg tgagcgtaac
 300
 gccaccaga gcggcttctc caccggccgga gggcacttcc tggcgatgga agaccacccc
 360
 ggggctgccc atctgaagct ggggtgcgct ggcggaaca gcgtcttcac actgggcaat
 420
 ggcaaagtcg ccggcgcgca actgcgcacc aacgccccac atgcaattga catcgtcttc
 480
 gctcaaacac gaagtgcccg gcgtgtactc attgtcgatg ggcaccgggg acccgcgggc
 540
 cg
 542

<210> 490

<211> 180

<212> PRT

<213> Homo sapiens

<400> 490

Xaa Ala Phe Gly Val Leu Ser Ala Val Val Asp Gly Asp Asp Ser Gly
 1 5 10 15
 Lys Pro Leu Leu Asn Gln His Gly Cys Tyr Lys Val Arg Phe Pro Phe
 20 25 30
 Thr Arg Asp Gln Lys Pro Ser Thr Arg Gly Ser Ala Trp Leu Arg Arg
 35 40 45
 Val Ser Leu Ser Ala Gly Ser Ser His Gly Met His Phe Pro Leu Leu
 50 55 60
 Lys Gly Ser Glu Val Leu Val Ser Phe Leu Gly Gly Asp Pro Asp Arg
 65 70 75 80
 Pro Ile Ile Val Gly Cys Val Pro Asn Ser Glu Thr Pro Ser Met Val
 85 90 95
 Val Glu Arg Asn Ala Thr Gln Ser Gly Phe Ser Thr Ala Gly Gly His
 100 105 110
 Phe Leu Ala Met Glu Asp His Pro Gly Ala Ala His Leu Lys Leu Gly
 115 120 125
 Ala Pro Gly Gly Asn Ser Val Phe Thr Leu Gly Asn Gly Lys Val Ala
 130 135 140
 Gly Ala Gln Leu Arg Thr Asn Ala Pro His Ala Ile Asp Ile Val Phe
 145 150 155 160
 Ala Gln Thr Arg Ser Ala Arg Arg Val Leu Ile Val Asp Gly His Arg
 165 170 175
 Gly Pro Gly Gly

180

<210> 491
 <211> 825
 <212> DNA
 <213> Homo sapiens

<400> 491
 nacgcgtcga ggcgacggtc ggcccgctca tggcgactgt tctcgagggc acatgggaac
 60
 gcatcgggtgc cggattccgg actgccttaa ccacagcctt ggaacgcacc gatgaatggg
 120
 tgggcggccc tgacagcaag cccctcaacg aagtcgagac actgcgccgg tgcgccgatg
 180
 aactcatcgg cgggcccgtc ggcgcggttg ccgcgatgca cggagggtca atcgaattgg
 240
 tgcagctgtc ggtcgggtgac gaagagcgca gactcgacgt caccatgaag ggagcatgcc
 300
 gaggttgcgc ggagccatc agaccctaca tcagcgctg gaacatcaac tgagtctgcg
 360
 nattgcgcga gccggtcacc gtgcgggaaa tctgacacct actccgacag ctccacctcg
 420
 acgagcacct ccacgacgag gccaaagccac tcgtagacgc attcctctctc ggcaccaat
 480
 tcctcccggg ccgcccggagc gacttcgtcg gcagtaacct ggtcgatgat ccctagcctg
 540
 gcggccatca tgccacgcag cgcattgaca gtacgaagcc aacgttgctg catcacaggg
 600
 ttcatggaga tacagccggt tcggtgcaac gtctccacat cagcacttaa ggactgagcg
 660
 tcttcccagc gcgcccgac atcctcggcg tcatggtcga catggaattg cgcgtcagct
 720
 gactcgtcgt cacgataggc gctgggcagg atcaatcgac gcacctcgtc gtcctcctgg
 780
 agtccagaaa actggtctct ccaaaaagcg aacgggtccc cctcc
 825

<210> 492
 <211> 58
 <212> PRT
 <213> Homo sapiens

<400> 492
 Met Asn Gly Trp Ala Ala Leu Thr Ala Ser Pro Ser Thr Lys Ser Arg
 1 5 10 15
 His Cys Ala Gly Ala Pro Met Asn Ser Ser Ala Gly Pro Ser Ala Arg
 20 25 30
 Leu Pro Arg Cys Thr Glu Gly Gln Ser Asn Trp Ser Thr Cys Arg Ser
 35 40 45
 Val Thr Lys Ser Ala Glu Ser Thr Ser Pro
 50 55

<210> 493
 <211> 863

<212> DNA

<213> Homo sapiens

<400> 493

nacgcgttcc aacctcgtca aaacggctat cgcaggaaat gaccccaact ggggtcgcac
 60
 cctcgcggcg atcggatgtg ttcctgagaa tatagctccc ttcgatcccg accaggtgga
 120
 tgtgtccatc aatgacattc agatctgtaa ggccgggggt atcggggagg accgcaacct
 180
 cgctgatatg aggccacgag aggttcacat cgatattgag ctgcatgcgg gtgatgccga
 240
 agctgcggta tggactaatg atctgacca ccaatacgtc gaagagaata gcgcgtatac
 300
 atcatgaccc ttgctcttga catccccctc aacgactccc agttctcggc tcagcggaaa
 360
 tctgaggtcc tggtagaagc gctgccttgg atcaggcggt ttcagggccg cactgtcgtc
 420
 gtgaaatag gcggaacgc gatggttgat cccggtctgc agcaggcctt cgccgacgac
 480
 attgtgttta tggcctctgt ggggattcgc cctattgtcg tccacggtgg tggccctcag
 540
 atcaatgccca tgcttgctga atccgctacc ccggtggagt tccgtaatgg tttgcgggtg
 600
 acatctccgg aggtcatgga ggttgtccgg atggtgctcg tcgggcaggc gggccgtcag
 660
 ctcgtaacc gaatcaacgc ctatgcgccg ctagcagctg gcatgtcagg cgaggacttt
 720
 ggcctttttt cggcccgga gtcgcgggta attgttgatg gcgagcaa atagacatgggt
 780
 ttagtgggag acatcgttga cgtcaacatc gatctcgta tctctatgct tgatcgcggt
 840
 cagattccgg tcattgcacc ggt
 863

<210> 494

<211> 186

<212> PRT

<213> Homo sapiens

<400> 494

Met Thr Leu Ala Leu Asp Ile Pro Leu Asn Asp Ser Gln Phe Ser Ala
 1 5 10 15
 Gln Arg Lys Ser Glu Val Leu Val Glu Ala Leu Pro Trp Ile Arg Arg
 20 25 30
 Phe Gln Gly Arg Thr Val Val Val Lys Tyr Gly Gly Asn Ala Met Val
 35 40 45
 Asp Pro Gly Leu Gln Gln Ala Phe Ala Asp Asp Ile Val Phe Met Ala
 50 55 60
 Ser Val Gly Ile Arg Pro Ile Val Val His Gly Gly Gly Pro Gln Ile
 65 70 75 80
 Asn Ala Met Leu Ala Glu Ser Ala Thr Pro Val Glu Phe Arg Asn Gly
 85 90 95
 Leu Arg Val Thr Ser Pro Glu Val Met Glu Val Val Arg Met Val Leu

```

      100      105      110
Val Gly Gln Val Gly Arg Gln Leu Val Asn Arg Ile Asn Ala Tyr Ala
      115      120      125
Pro Leu Ala Ala Gly Met Ser Gly Glu Asp Phe Gly Leu Phe Ser Ala
      130      135      140
Arg Lys Ser Arg Val Ile Val Asp Gly Glu Gln Ile Asp Met Gly Leu
145      150      155      160
Val Gly Asp Ile Val Asp Val Asn Ile Asp Leu Val Ile Ser Met Leu
      165      170      175
Asp Arg Gly Gln Ile Pro Val Ile Ala Pro
      180      185

```

<210> 495
 <211> 514
 <212> DNA
 <213> Homo sapiens

```

<400> 495
gcgcgcgaca ccggtgcccc gattagcgtg ccagtgggtg acgtcactaa gggtcacgtc
60
tggaatgtga caggtgacgt tcttaacgcc ngatccctcc acaatcgagg tgacnntgag
120
cgttggccga tccaccggga tccccgggcc ttgatgacc ttgagcccgga gaccgagatg
180
ctggagaccg gtattaaggt ccttgacttg ctgactcctt acgtcaaggg cggcaagatt
240
ggcctctttg gcggcgctgg tgtgggtaag acggtgctca ttcaggagat gatttaccgt
300
atcgcccaca acttcggcgg tacttcggtt ttgcgcggtg tcggtgagcg taccgcgag
360
ggtaacgacc tcatcaacga gatggacgag gccggtgtgc tcaaagacac cgccctggta
420
ttcggccaga tggacgagcc cccgggcacg cggtacgagc tgtcgcgctg gcagccctgc
480
ggcccatgcc tggtaactg ctgtgggacc ttgg
514

```

<210> 496
 <211> 171
 <212> PRT
 <213> Homo sapiens

```

<400> 496
Ala Arg Asp Thr Gly Ala Pro Ile Ser Val Pro Val Gly Asp Val Thr
1      5      10      15
Lys Gly His Val Trp Asn Val Thr Gly Asp Val Leu Asn Ala Xaa Ser
      20      25      30
Leu His Asn Arg Gly Asp Xaa Glu Arg Trp Pro Ile His Arg Asp Pro
      35      40      45
Pro Ala Phe Asp Asp Leu Glu Pro Glu Thr Glu Met Leu Glu Thr Gly
      50      55      60
Ile Lys Val Leu Asp Leu Leu Thr Pro Tyr Val Lys Gly Gly Lys Ile
65      70      75      80
Gly Leu Phe Gly Gly Ala Gly Val Gly Lys Thr Val Leu Ile Gln Glu

```

```

<400> 498
Met Asn Glu Glu Lys Thr Gln Pro His Lys Arg Asp Thr Arg Trp Arg
 1             5             10             15
Gly Glu Thr Gln Ser Ser Phe Leu Ser Val Asp Ser Glu Gln Arg
      20             25             30
Arg Gly Ala Pro Ser Phe Val Phe Ser Ser Ser Gly Glu Arg Met Asp

```



```

      35          40          45
Cys Leu His Ala Ser Cys His Thr Pro Ala Val Ile Pro Ala Arg Ala
 50          55          60
Pro Ser Ala Glu Ala Glu Leu Cys Ser Ala Gln Ala Trp Asp Leu Pro
65          70          75          80
Arg Gln Ala Pro Val Gly Gly Ala Ala Pro Gly Lys Glu Ala Thr Ala
      85          90          95
Ser Leu Asn Ile Leu Arg Cys Lys Val Val Ala Pro Arg Gly Val Ser
      100          105          110
Val Lys Thr Gly Thr Arg Met Ala Gly Pro Ala Arg Leu Phe Pro His
      115          120          125
Leu Ser Ala Ser Glu Ala Ser Leu Glu Asp Ser Gly Pro Arg Met Ser
      130          135          140
Pro Arg Thr Ser Gln Ser Ala Ser Ser Ser Tyr Phe Cys Cys Ser Leu
145          150          155          160
Gly Pro Asp Leu Ala Lys Val Ser Gln Arg Gly Gly Pro Arg Ser Glu
      165          170          175
Leu Ser Ser Cys Arg Gly Pro Arg Asp Gly Leu Gly Cys Lys Leu
      180          185          190

```

<210> 499
 <211> 444
 <212> DNA
 <213> Homo sapiens

```

<400> 499
acgcgtgaag ggtgggcagt gttgagctga gtgagccctc ctccctgcaa tgctggagcc
60
ctgccttctg cctgaccctc tggcttccta agcagtctat acgtgagaag ccctttcttc
120
aagtgaagc ttctgagctc actacgagag cactggagct ggaacctctc tgggttcaaa
180
tcctcaactg gggggttga ggaggttact tcacttctca aaacctcaat ttccttatct
240
gcaaaatggg gtaataggag cccctcttca tcaatgcttg gagggaatgc ctggcacagt
300
agggcagtta ccgtcatgga gaacagaaag gccccgagct atcctggatg tggtgagaat
360
gggtcctgga tcctgcctgc tcggcctttt cattctcttc ttcacctaca ggctccaca
420
aagggcctct gaaaacacag ggtg
444

```

<210> 500
 <211> 105
 <212> PRT
 <213> Homo sapiens

```

<400> 500
Met Thr Val Thr Ala Leu Leu Cys Gln Ala Phe Pro Pro Ser Ile Asp
 1          5          10          15
Glu Glu Gly Leu Leu Leu Pro His Phe Ala Asp Lys Glu Ile Glu Val
      20          25          30
Leu Arg Ser Glu Val Thr Ser Ser Asn Pro Pro Val Glu Asp Leu Asn

```

```

      35          40          45
Pro Glu Arg Phe Gln Leu Gln Cys Ser Arg Ser Glu Leu Arg Ser Phe
   50          55          60
His Leu Lys Lys Gly Leu Thr Tyr Arg Leu Leu Arg Lys Pro Glu
65          70          75          80
Gly Gln Ala Glu Gly Arg Ala Pro Ala Leu Gln Gly Gly Gly Leu Thr
      85          90          95
Gln Leu Asn Thr Ala His Pro Ser Arg
      100          105

```

<210> 501
 <211> 800
 <212> DNA
 <213> Homo sapiens

```

<400> 501
agatctgac cgagaagtgg ctgctcaggg aaatgactac tccatgggtt tcttaactca
60
ggtactcctt attcaatgag aggcctgagg tgagaccgc catgcggcgc gtggatcgca
120
tggtgttagt gcacactagc aaggggctta ggtctccagc tgaggtcaga tgcacacttg
180
gaccttgtag tggggagtaa cacacatctc tgtgttcagc gaaccatcca ggagctgttt
240
gaagtttatt ctcccatgga tgatgctggc ttcccggcca aagctgagga gtttgtggtg
300
ctttctcagg aaccttctgt cacggaaacc attgcacca aaattgcaag acctttcata
360
gaggccctca agagtattga gtatctggag gaggatgccc agaagtccgc acaggagggg
420
gtgctgggac cacacactga tgctctgtca tcagactctg agaacatgcc gtgtgatgaa
480
gaaccatccc aattagagga gctagctgac ttcattggagc agcttacacc aattgaaaaa
540
tatgctttaa attacctgga atcttgaggc agggcctgag agagcacgct gcgccgtact
600
tccagcagct gcggcagacc acggctccac gcctgtgca gttccctgag ctgaggctgg
660
tgcagttcga ctcaggtagt cggcagttgg gggcgtggcc cgtgcgggag ctgcactggc
720
cctggatgat gaggcgctct tgatgtgatt cgtttcccag ggaagttgga agcttttagct
780
atcttgcttc agaaactgaa
800

```

<210> 502
 <211> 103
 <212> PRT
 <213> Homo sapiens

```

<400> 502
Met Asp Asp Ala Gly Phe Pro Val Lys Ala Glu Glu Phe Val Val Leu
 1          5          10          15
Ser Gln Glu Pro Ser Val Thr Glu Thr Ile Ala Pro Lys Ile Ala Arg

```

```

      20      25      30
Pro Phe Ile Glu Ala Leu Lys Ser Ile Glu Tyr Leu Glu Glu Asp Ala
      35      40      45
Gln Lys Ser Ala Gln Glu Gly Val Leu Gly Pro His Thr Asp Ala Leu
      50      55      60
Ser Ser Asp Ser Glu Asn Met Pro Cys Asp Glu Glu Pro Ser Gln Leu
65      70      75      80
Glu Glu Leu Ala Asp Phe Met Glu Gln Leu Thr Pro Ile Glu Lys Tyr
      85      90      95
Ala Leu Asn Tyr Leu Glu Ser
      100

```

<210> 503
 <211> 538
 <212> DNA
 <213> Homo sapiens

```

<400> 503
nnacgcgttg tcgtctctcc gatcattgat tttgttgat tctgcaatga tgtaaaggaa
60
gatgatgaca cggagaagtt taaagaagcc attgtgaaat ttcataggct gtttgggatg
120
ccagaggaag agaaactcgt caactattac tcttgcagct attggaaggg gaagggtcccc
180
cgtcagggtt ggatgtacct cagcattaac cacctttgct tttattcttt tcttatggga
240
agggaagcga aactgggtcat ccggtgggta gacatcactc agcttgagaa gaatgcccc
300
ctgcttctgc ctgatgtgat caaagtgagc acacggtcca gtgagcattt cttctctgta
360
ttcctcaaca tcaacgagac cttcaagtta atggagcagc ttgccaacat agccatgagg
420
caactcttag acaatgaggg atttgaacaa gatcgatccc tgcccaaact caaaaggaaa
480
tctcctaaaa aagtgtctgc tctaaaacgt gatcttgatg cctgggccct tcacgcgt
538

```

<210> 504
 <211> 179
 <212> PRT
 <213> Homo sapiens

```

<400> 504
Xaa Arg Val Val Val Ser Pro Ile Ile Asp Phe Val Val Phe Cys Asn
1      5      10      15
Asp Val Lys Glu Asp Asp Asp Thr Glu Lys Phe Lys Glu Ala Ile Val
      20      25      30
Lys Phe His Arg Leu Phe Gly Met Pro Glu Glu Glu Lys Leu Val Asn
      35      40      45
Tyr Tyr Ser Cys Ser Tyr Trp Lys Gly Lys Val Pro Arg Gln Gly Trp
      50      55      60
Met Tyr Leu Ser Ile Asn His Leu Cys Phe Tyr Ser Phe Leu Met Gly
65      70      75      80
Arg Glu Ala Lys Leu Val Ile Arg Trp Val Asp Ile Thr Gln Leu Glu

```

```

      85              90              95
Lys Asn Ala Pro Leu Leu Leu Pro Asp Val Ile Lys Val Ser Thr Arg
      100              105              110
Ser Ser Glu His Phe Phe Ser Val Phe Leu Asn Ile Asn Glu Thr Phe
      115              120              125
Lys Leu Met Glu Gln Leu Ala Asn Ile Ala Met Arg Gln Leu Leu Asp
      130              135              140
Asn Glu Gly Phe Glu Gln Asp Arg Ser Leu Pro Lys Leu Lys Arg Lys
      145              150              155              160
Ser Pro Lys Lys Val Ser Ala Leu Lys Arg Asp Leu Asp Ala Trp Ala
      165              170              175
Leu His Ala

```

<210> 505
 <211> 381
 <212> DNA
 <213> Homo sapiens

```

<400> 505
gtgcacgaca ccgaacggta cgaacgtatc tcccaggcac gtcgcgagga acagcaggcc
60
atgctcggct acgaacgctc aagaacctgt cgcattgacct tgctcacggg gcagctggac
120
gacccctcca cgactccttg cggacgctgc gacgtctgtg ctggcccgtg gtactcagtc
180
gaggtcgatc agtcagccgc tgtgagagcc gtccaatccc tcaaccgggt gggagttccg
240
gtggaaccac gcgcgcctg gcccgagggt atggacgccc tccaggttgc gctcaagggt
300
cgcatcagtg ccgaggagat cgctgcagag ggccgcgtca tcgccagact ctccgatctg
360
ggttggggag gggcgctgcg c
381

```

<210> 506
 <211> 127
 <212> PRT
 <213> Homo sapiens

```

<400> 506
Val His Asp Thr Glu Arg Tyr Glu Arg Ile Ser Gln Ala Arg Arg Glu
1      5      10      15
Glu Gln Gln Ala Met Leu Gly Tyr Asp Xaa Ser Arg Thr Cys Arg Met
20      25      30
Thr Leu Leu Thr Gly Gln Leu Asp Asp Pro Ser Thr Thr Pro Cys Gly
35      40      45
Arg Cys Asp Val Cys Ala Gly Pro Trp Tyr Ser Val Glu Val Asp Gln
50      55      60
Ser Ala Ala Val Arg Ala Val Gln Ser Leu Asn Arg Val Gly Val Pro
65      70      75      80
Val Glu Pro Arg Ala Ala Trp Pro Ala Gly Met Asp Ala Leu Gln Val
85      90      95
Ala Leu Lys Gly Arg Ile Ser Ala Glu Glu Ile Ala Ala Glu Gly Arg

```

| | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| | 100 | | 105 | | 110 | | | | | | | | | |
| Val | Ile | Ala | Arg | Leu | Ser | Asp | Leu | Gly | Trp | Gly | Gly | Ala | Leu | Arg |
| | 115 | | | | 120 | | | | | | | 125 | | |

<210> 507
 <211> 499
 <212> DNA
 <213> Homo sapiens

<400> 507
 gccggcggtgt tcaacctcat ggtgtgggcc ttcattaccg acgtcatcga tgcccaggag
 60
 gtcattgtccg gggagcgtga agacggtgtc atctatggcg tgaactcctt cgcccgcaaa
 120
 cttgcccagg ccattgccgg tggaatcggc ggagccatgc tgacgatgat cggctaccag
 180
 tcttctctccc aaggtggtgc cgttcagtcg gagtccgtcg tcaatcacct gtacacgctc
 240
 gccaccgcca tcccgacgat ctgctgcctc ggcgctgccc tgctcatgct gggctacccg
 300
 ctcacccgcg acaaggtggt cgccaacgcc gacgagttgg ctgctcgcca cgcagtacag
 360
 gccgagcaaa actcctgacc cataacggag gcacatcatg gacacgctca tgcggatcac
 420
 cgaccacttg acaacctcgc cgggtatcca attgaaaatt gacaagcgat ggggtgcctc
 480
 cgtcacattt gtgacgcgt
 499

<210> 508
 <211> 125
 <212> PRT
 <213> Homo sapiens

<400> 508
 Ala Gly Val Phe Asn Leu Met Val Trp Ala Phe Ile Thr Asp Val Ile
 1 5 10 15
 Asp Ala Gln Glu Val Met Ser Gly Glu Arg Glu Asp Gly Val Ile Tyr
 20 25 30
 Gly Val Asn Ser Phe Ala Arg Lys Leu Ala Gln Ala Ile Ala Gly Gly
 35 40 45
 Ile Gly Gly Ala Met Leu Thr Met Ile Gly Tyr Gln Ser Ser Ser Gln
 50 55 60
 Gly Gly Ala Val Gln Ser Glu Ser Val Val Asn His Leu Tyr Thr Leu
 65 70 75 80
 Ala Thr Ala Ile Pro Thr Ile Cys Cys Leu Gly Ala Ala Leu Leu Met
 85 90 95
 Leu Gly Tyr Pro Leu Thr Arg Asp Lys Val Val Ala Asn Ala Asp Glu
 100 105 110
 Leu Ala Arg Arg His Ala Val Gln Ala Glu Gln Asn Ser
 115 120 125

<210> 509
 <211> 360

<212> DNA

<213> Homo sapiens

<400> 509

ttggccatgg atttggtcgc caagttcagt cccaaagatg tcacgctcta tctaattggac
 60
 ttcgggacca atgggtgtggc accactaggc caattaccac aggtggccga caccttgctt
 120
 ttggatcata cggagaagat tgccaagttt gtacgcatca tggagcggga gctcaaccgg
 180
 cgtaagaagc tcttgtccga ctacggtggt ggtacactag agctctaccg tcaggctagc
 240
 ggtcagcaag agccggccat cgtcatcctg ctggacagtt atgagtccat gaaggaagag
 300
 gcctatgaag cggagctctt cagcgtcttg gtgcggatct cccgggaagg tctcagcatc
 360

<210> 510

<211> 120

<212> PRT

<213> Homo sapiens

<400> 510

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Leu | Ala | Met | Asp | Leu | Ala | Arg | Lys | Phe | Ser | Pro | Lys | Asp | Val | Thr | Leu |
| 1 | | | 5 | | | | | 10 | | | | | 15 | | |
| Tyr | Leu | Met | Asp | Phe | Gly | Thr | Asn | Gly | Val | Ala | Pro | Leu | Gly | Gln | Leu |
| | | | 20 | | | | 25 | | | | | 30 | | | |
| Pro | Gln | Val | Ala | Asp | Thr | Leu | Leu | Leu | Asp | His | Thr | Glu | Lys | Ile | Ala |
| | | 35 | | | | 40 | | | | | 45 | | | | |
| Lys | Phe | Val | Arg | Ile | Met | Glu | Arg | Glu | Leu | Asn | Arg | Arg | Lys | Lys | Leu |
| | 50 | | | | 55 | | | | | 60 | | | | | |
| Leu | Ser | Asp | Tyr | Gly | Val | Gly | Thr | Leu | Glu | Leu | Tyr | Arg | Gln | Ala | Ser |
| 65 | | | | 70 | | | | | 75 | | | | | 80 | |
| Gly | Gln | Gln | Glu | Pro | Ala | Ile | Val | Ile | Leu | Leu | Asp | Ser | Tyr | Glu | Ser |
| | | | 85 | | | | 90 | | | | | | 95 | | |
| Met | Lys | Glu | Glu | Ala | Tyr | Glu | Ala | Glu | Leu | Phe | Thr | Leu | Leu | Val | Arg |
| | | 100 | | | | | | 105 | | | | | 110 | | |
| Ile | Ser | Arg | Glu | Gly | Leu | Ser | Ile | | | | | | | | |
| | | 115 | | | | 120 | | | | | | | | | |

<210> 511

<211> 361

<212> DNA

<213> Homo sapiens

<400> 511

ntcgcgaacc gcggctatgc ggtgctccag cccaatttcc gcggatcggg cggttatggc
 60
 actgcgttcg gcgatgccgg catcggccag atcgggcgca agatgcagga cgatctcgac
 120
 gacgggatgg actggctggt caaggagggc atcgtcgaca agggccgggt gtgcatcgtc
 180
 ggggcctcct atggcggcta tgccgcgatg tggggcgaga tccgcaatcc cgaacgctat
 240

cgctgcgcgg cgagcctggc gggggttgcc gattaaggcc atgctcaaatt ataaccggcg
 300
 ctatctcgac aaggaggcgg gcaagcgctg gccgccccgn tcaaccggcg aaccggaatt
 360
 c
 361

<210> 512
 <211> 91
 <212> PRT
 <213> Homo sapiens

<400> 512
 Xaa Ala Asn Arg Gly Tyr Ala Val Leu Gln Pro Asn Phe Arg Gly Ser
 1 5 10 15
 Gly Gly Tyr Gly Thr Ala Phe Gly Asp Ala Gly Ile Gly Gln Ile Gly
 20 25 30
 Arg Lys Met Gln Asp Asp Leu Asp Asp Gly Met Asp Trp Leu Val Lys
 35 40 45
 Glu Gly Ile Val Asp Lys Gly Arg Val Cys Ile Val Gly Ala Ser Tyr
 50 55 60
 Gly Gly Tyr Ala Ala Met Trp Gly Ala Ile Arg Asn Pro Glu Arg Tyr
 65 70 75 80
 Arg Cys Ala Ala Ser Leu Ala Gly Val Ala Asp
 85 90

<210> 513
 <211> 369
 <212> DNA
 <213> Homo sapiens

<400> 513
 nnatgcagac tagaagatgg catgacgggt ttggctggcg gtttcgggct atgcggcatt
 60
 ccagaaaatc tgattcaaga gatcaaacga cgccagactt gtgatttgac catagtgtca
 120
 aataactgtg gtgtagatgg ttttggttta ggggttttgc tagaagataa gcaagtacgc
 180
 aaaatggtgt cttcttatgt gggtgaaaat gcaactgtttg agaagcaatt attacaaggt
 240
 gagttggaag tcgagctcac tcttcaaggc actcttgccg aaaaactacg cgctggcggc
 300
 gcgggaattc ctgccttttt cacagcaacg ggtgtaggta cacctattgg tgagggtaaa
 360
 gacacgcgt
 369

<210> 514
 <211> 123
 <212> PRT
 <213> Homo sapiens

<400> 514
 Xaa Cys Arg Leu Glu Asp Gly Met Thr Val Leu Ala Gly Gly Phe Gly

```

      1             5             10             15
Leu Cys Gly Ile Pro Glu Asn Leu Ile Gln Glu Ile Lys Arg Arg Gln
      20             25             30
Thr Cys Asp Leu Thr Ile Val Ser Asn Asn Cys Gly Val Asp Gly Phe
      35             40             45
Gly Leu Gly Val Leu Leu Glu Asp Lys Gln Val Arg Lys Met Val Ser
      50             55             60
Ser Tyr Val Gly Glu Asn Ala Leu Phe Glu Lys Gln Leu Leu Gln Gly
      65             70             75             80
Glu Leu Glu Val Glu Leu Thr Pro Gln Gly Thr Leu Ala Glu Lys Leu
      85             90             95
Arg Ala Gly Gly Ala Gly Ile Pro Ala Phe Phe Thr Ala Thr Gly Val
      100            105            110
Gly Thr Pro Ile Gly Glu Gly Lys Asp Thr Arg
      115            120

```

<210> 515
 <211> 387
 <212> DNA
 <213> Homo sapiens

```

<400> 515
gcgtagggacg agaaggccgc cggcaactgc gcgatcgact acgggttcca ccagatcctc
60
tccgacgtgc aggactcgtc gctgaccgcg atggacgagc tgatcaccga gggcgtgaca
120
tccttcaagc tcttcgtggc ctacaagggc gtcttcctct cggacgacgg gcagatcctg
180
cgggcggttc agaaggccgc cgacaacggc gcgatgatga tgatgcacgc cgagaacggc
240
gcatcatcg acgtgctcgt gcagcaggcg ctcgaggccg ggaagaccac cccgtactac
300
cacggcatca gccggccgtg gcaggccgag gaggaggcca cccaccgccc gatcatgatc
360
gccgacctga ccggtgcgcc gttgtac
387

```

<210> 516
 <211> 129
 <212> PRT
 <213> Homo sapiens

```

<400> 516
Ala Trp Asp Glu Lys Ala Ala Gly Asn Cys Ala Ile Asp Tyr Gly Phe
1             5             10             15
His Gln Ile Leu Ser Asp Val Gln Asp Ser Ser Leu Thr Ala Met Asp
      20             25             30
Glu Leu Ile Thr Glu Gly Val Thr Ser Phe Lys Leu Phe Val Ala Tyr
      35             40             45
Lys Gly Val Phe Leu Ser Asp Asp Gly Gln Ile Leu Arg Ala Phe Gln
      50             55             60
Lys Gly Ala Asp Asn Gly Ala Met Met Met Met His Ala Glu Asn Gly
      65             70             75             80
Ala Ile Ile Asp Val Leu Val Gln Gln Ala Leu Glu Ala Gly Lys Thr

```



```

      85              90              95
Thr Pro Tyr Tyr His Gly Ile Ser Arg Pro Trp Gln Ala Glu Glu Glu
      100              105              110
Ala Thr His Arg Ala Ile Met Ile Ala Asp Leu Thr Gly Ala Pro Leu
      115              120              125
Tyr

```

<210> 517
 <211> 377
 <212> DNA
 <213> Homo sapiens

```

<400> 517
acgcgtgaag ggctggtggg caggccttgc gccccctctg gggacagctc tcctccaccc
60
agacccttc gggccaacag tggggagggg ctgccgtctg agccactgtt ccgacagggg
120
attcgcgagt tccgggggag ctggggactg agctgcgggc ctctctgggct ggggctcttc
180
tccgaggttg gaggcagctt tagaaacttg agacccttag ctggagaggg cagaaggggt
240
ccctgagctt cccagggaga agggggggcca atttgagct tgcttttcac ctgagatgag
300
gaatgggggt ggccaggccg agagcccagt ggggcatccc cagcacccat gaacatgcta
360
aggaagggga ggggccc
377

```

<210> 518
 <211> 118
 <212> PRT
 <213> Homo sapiens

```

<400> 518
Met Phe Met Gly Ala Gly Asp Ala Pro Leu Gly Ser Arg Pro Gly His
1      5      10      15
Pro His Ser Ser Ser Gln Val Lys Ser Lys Leu Gln Ile Gly Pro Pro
20     25     30
Ser Pro Gly Glu Ala Gln Gly Pro Leu Leu Pro Ser Pro Ala Arg Gly
35     40     45
Leu Lys Phe Leu Lys Leu Pro Thr Ser Glu Lys Ser Pro Ser Pro
50     55     60
Gly Gly Pro Gln Leu Ser Pro Gln Leu Pro Arg Asn Ser Arg Ile Pro
65     70     75     80
Cys Arg Asn Ser Gly Ser Asp Gly Ser Pro Ser Pro Leu Leu Ala Arg
85     90     95
Arg Gly Leu Gly Gly Gly Glu Leu Ser Pro Glu Gly Ala Gln Gly Leu
100    105    110
Pro Thr Ser Pro Ser Arg
115

```

<210> 519
 <211> 311

<212> DNA

<213> Homo sapiens

<400> 519

g c g c g c c a g g g g g a a g g g a g a g a a a c a c a g a a a a t g a g g g g a a a t a c c a g a t a c t g a
 60
 a g a a t t t a a a t t a t t a t a a a g g a a c c t t t t c t g c a a c t c t g a a a a t g t t a g a a t a t c c a
 120
 a a g a a a t t g a t a a t t t t c t a g g a a a c a t g a c t t a c c a a a a t t a a c t c t a g a a a g a a t c
 180
 g a t a c a c a t c a g t a a c a a c a g a a g t t g a g a a g t a g t t a a c a t a t t g c c a a a c c t g g a a t
 240
 t c a t g a t t g a a t t c t t t g a g a t c t a c t g t g a g t a c a t a c t c t g c c t c t g t t c a g e t g t t c
 300
 c a g a a c t t a a g
 311

<210> 520

<211> 92

<212> PRT

<213> Homo sapiens

<400> 520

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Arg | Gly | Lys | Tyr | Gln | Ile | Leu | Lys | Asn | Leu | Asn | Tyr | Tyr | Lys | Gly |
| 1 | | | | 5 | | | | | 10 | | | | | 15 | |
| Thr | Phe | Ser | Ala | Thr | Leu | Lys | Asn | Val | Arg | Ile | Ser | Lys | Glu | Ile | Asp |
| | | | 20 | | | | | 25 | | | | | 30 | | |
| Asn | Phe | Leu | Gly | Lys | His | Asp | Leu | Pro | Lys | Leu | Thr | Leu | Glu | Lys | Asn |
| | | | 35 | | | | 40 | | | | | 45 | | | |
| Arg | Tyr | Thr | Ser | Val | Thr | Thr | Glu | Val | Glu | Lys | Val | Val | Asn | Ile | Leu |
| | 50 | | | | | 55 | | | | | 60 | | | | |
| Pro | Asn | Leu | Glu | Phe | Met | Ile | Glu | Phe | Phe | Glu | Ile | Tyr | Cys | Glu | Tyr |
| 65 | | | | 70 | | | | | 75 | | | | | 80 | |
| Ile | Leu | Cys | Leu | Cys | Ser | Ala | Val | Pro | Glu | Leu | Lys | | | | |
| | | | 85 | | | | | | 90 | | | | | | |

<210> 521

<211> 352

<212> DNA

<213> Homo sapiens

<400> 521

n n n g a t g c c a c g c c g g t c t a c g g a a t c t c c a c c g g c t t c g g c g c g t t g c c g c c g c c a c
 60
 a t t c c a g a a g a g a t g c g c g c g c a g c t g c a g c t g t c c c t g g t g c g t c c c a c g c g g c c g g c
 120
 a c c g g c c c t g a g g t g g a a g a a g a g t a a t t c g c g c g c t c a t g c t g t g c g c c t a t c c a c c
 180
 c t g t g t a c c g g c c g t a c c g g c g t g c g c c c c g t g g t g g t a g a a a c t t a t g c c a a g g c g c t c
 240
 a a c g c c g g c a t c g t g c c g g g g g t g c g c g a a t a c g g g t c g c t g g g c t g c t c g g c g a c t t g
 300
 g c c c c g c t g g c t c a c t g c g c c c t a g c g c t g t t g g g t g a g g g t g a g g t a c g c n
 352

<210> 522
 <211> 117
 <212> PRT
 <213> Homo sapiens

<400> 522
 Xaa Asp Ala Thr Pro Val Tyr Gly Ile Ser Thr Gly Phe Gly Ala Leu
 1 5 10 15
 Ala Arg Arg His Ile Pro Glu Glu Met Arg Ala Gln Leu Gln Leu Ser
 20 25 30
 Leu Val Arg Ser His Ala Ala Gly Thr Gly Pro Glu Val Glu Glu Glu
 35 40 45
 Val Ile Arg Ala Leu Met Leu Leu Arg Leu Ser Thr Leu Cys Thr Gly
 50 55 60
 Arg Thr Gly Val Arg Pro Val Val Val Glu Thr Tyr Ala Lys Ala Leu
 65 70 75 80
 Asn Ala Gly Ile Val Pro Gly Val Arg Glu Tyr Gly Ser Leu Gly Cys
 85 90 95
 Ser Gly Asp Leu Ala Pro Leu Ala His Cys Ala Leu Ala Leu Leu Gly
 100 105 110
 Glu Gly Glu Val Arg
 115

<210> 523
 <211> 693
 <212> DNA
 <213> Homo sapiens

<400> 523
 agcgcttcca cagtcgcgca aactcctctt ggtctagccg cccattcact ttcagttcca
 60
 tcagagccac caagctgcgg caccatctaa ggagaacatg tcccctggag gtcctgttag
 120
 aagctcctgg ttgagaaggc cctgaagctg ggtggcatca atgtccagcc tctgctgagc
 180
 atatctgttg aaaatgcttt gttgggagcc atgttctgaa gggcttcctt tcattctgag
 240
 gttgaaatgg ctgtcaggt gcctgtcact gtctggcatt ttcaggaaga ttcggagcaa
 300
 gaactccgct gattttctcc gtgtctgtgc aaccacaaca tagttcccag ggctcagatg
 360
 gtaagtcag gtgaagttgc ggcggaattt attatttgag ctttggacag tgtttctgaa
 420
 cgaggaaaaa aacacgggtg gaaatttctc ccggaaccgc tgtgagccag ccagaatcac
 480
 ttggaatcg agtggaatt ttgcatcttc tgctttcaaa tttgatggtg tgacagcaac
 540
 tgtgacgcac acgacaacat tggcgcttc cattggctct tgcacagaga agttgaattg
 600
 agcatcattt ccgggtcttc ctggcggtgtt tcctagaatc attgcttcct aaacattatt
 660
 tgggaccatc cttcgtggag tgtgtttcca tgg
 693

<210> 524
 <211> 193
 <212> PRT
 <213> Homo sapiens

<400> 524
 Met Ile Leu Gly Asn Thr Pro Gly Gly Pro Gly Asn Asp Ala Gln Phe
 1 5 10 15
 Asn Phe Ser Val Gln Glu Pro Met Glu Gly Thr Asn Val Val Val Cys
 20 25 30
 Val Thr Val Ala Val Thr Pro Ser Asn Leu Lys Ala Glu Asp Ala Lys
 35 40 45
 Phe Pro Leu Asp Phe Gln Val Ile Leu Ala Gly Ser Gln Arg Phe Arg
 50 55 60
 Glu Lys Phe Pro Pro Val Phe Phe Ser Ser Phe Arg Asn Thr Val Gln
 65 70 75 80
 Ser Ser Asn Asn Lys Phe Arg Arg Asn Phe Thr Met Thr Tyr His Leu
 85 90 95
 Ser Pro Gly Asn Tyr Val Val Val Ala Gln Thr Arg Arg Lys Ser Ala
 100 105 110
 Glu Phe Leu Leu Arg Ile Phe Leu Lys Met Pro Asp Ser Asp Arg His
 115 120 125
 Leu Ser Ser His Phe Asn Leu Arg Met Lys Gly Ser Pro Ser Glu His
 130 135 140
 Gly Ser Gln Gln Ser Ile Phe Asn Arg Tyr Ala Gln Gln Arg Leu Asp
 145 150 155 160
 Ile Asp Ala Thr Gln Leu Gln Gly Leu Leu Asn Gln Glu Leu Leu Thr
 165 170 175
 Gly Pro Pro Gly Asp Met Phe Ser Leu Asp Gly Ala Ala Ala Trp Trp
 180 185 190
 Leu

<210> 525
 <211> 1101
 <212> DNA
 <213> Homo sapiens

<400> 525
 nggcaagttg caaagagagc ctcagaggtc cgaagagcgc tgcgctccta ctgcggttcg
 60
 cttcttcctc ttctcggttc cctactgtga aatcgagcgc acatttaca aggctccgg
 120
 gtcctaccga gaccgatccg cagcggttgg cccggtcgcg cctattgcat cgggagcccc
 180
 cgagcaccgg cgaaggactg gcgggtgggg tagggagggtg gcggcgccgg catggcgagg
 240
 ttcccgaagg cgcacctggc cgctgcagga gttatgttac ttgcccactt cttcacggac
 300
 cagtttcagt tcgccgatgg gaaacccgga gaccaaatac ttgattggca gtatggagtt
 360
 actcaggcct tcctcacac agaggaggag gtggaagttg attcacacgc gtacagccac
 420

aggtggaaaa gaaacttgga ctttctcaag gcggtagaca cgaaccgagc aagcgtcggc
 480
 caagactctc ttgagccag aagcttcaca gacctgctgc tggatgatgg gcaggacaat
 540
 aacactcaga tcgaggagga tacagaccac aattactata tatctcgaat atatggtcca
 600
 tctgattctg ccagccggga tttatgggtg aacatagacc aaatggaaaa agataaagtg
 660
 aagattcatg gaattattgtc caatactcat cggcaagctg caagagtga tctgtccttc
 720
 gattttccat tttatggcca ctctctacgt gaaatcactg tggcaaccgg gggtttcata
 780
 tacattggag aagtcgtaca tcgaatgcta acagccacac agtacatagc acctttaatg
 840
 gcaaatttcg atcccagtgt atccagaaat tcaactgtca gatattttga taatggcaca
 900
 gcacttgtgg tccagtggga ccatgtacat ctccaggata attataacct gggaagcttc
 960
 acattccagg caaccctgct catggatgga cgaatcatct ttggatacaa agaaattcct
 1020
 gtcttgggtca cacagataag ttcaaccaat catccagtga aagtcggact gtccgatgca
 1080
 tttgtcgttg tccacaggat c
 1101

<210> 526

<211> 290

<212> PRT

<213> Homo sapiens

<400> 526

| | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Ala | Arg | Phe | Pro | Lys | Ala | Asp | Leu | Ala | Ala | Gly | Val | Met | Leu |
| 1 | | | | 5 | | | | 10 | | | | | 15 | |
| Leu | Cys | His | Phe | Phe | Thr | Asp | Gln | Phe | Gln | Phe | Ala | Asp | Gly | Lys |
| | | | 20 | | | | 25 | | | | | 30 | | Pro |
| Gly | Asp | Gln | Ile | Leu | Asp | Trp | Gln | Tyr | Gly | Val | Thr | Gln | Ala | Phe |
| | | 35 | | | | | 40 | | | | | 45 | | Pro |
| His | Thr | Glu | Glu | Glu | Val | Glu | Val | Asp | Ser | His | Ala | Tyr | Ser | His |
| | | 50 | | | | 55 | | | | | 60 | | | Arg |
| Trp | Lys | Arg | Asn | Leu | Asp | Phe | Leu | Lys | Ala | Val | Asp | Thr | Asn | Arg |
| 65 | | | | 70 | | | | 75 | | | | | 80 | Ala |
| Ser | Val | Gly | Gln | Asp | Ser | Leu | Glu | Pro | Arg | Ser | Phe | Thr | Asp | Leu |
| | | | 85 | | | | | 90 | | | | | 95 | Leu |
| Leu | Asp | Asp | Gly | Gln | Asp | Asn | Asn | Thr | Gln | Ile | Glu | Glu | Asp | Thr |
| | | | 100 | | | | | 105 | | | | | 110 | Asp |
| His | Asn | Tyr | Tyr | Ile | Ser | Arg | Ile | Tyr | Gly | Pro | Ser | Asp | Ser | Ala |
| | | 115 | | | | 120 | | | | | | 125 | | Ser |
| Arg | Asp | Leu | Trp | Val | Asn | Ile | Asp | Gln | Met | Glu | Lys | Asp | Lys | Val |
| | | 130 | | | | 135 | | | | | 140 | | | Lys |
| Ile | His | Gly | Ile | Leu | Ser | Asn | Thr | His | Arg | Gln | Ala | Ala | Arg | Val |
| 145 | | | | 150 | | | | | | 155 | | | 160 | Asn |
| Leu | Ser | Phe | Asp | Phe | Pro | Phe | Tyr | Gly | His | Phe | Leu | Arg | Glu | Ile |
| | | | | 165 | | | | 170 | | | | | 175 | Thr |
| Val | Ala | Thr | Gly | Gly | Phe | Ile | Tyr | Thr | Gly | Glu | Val | Val | His | Arg |
| | | | | | | | | | | | | | | Met |

| | | | | | |
|---|-----|-----|-----|-----|-----|
| | 180 | | 185 | | 190 |
| Leu Thr Ala Thr Gln Tyr Ile Ala Pro Leu Met Ala Asn Phe Asp Pro | | | | | |
| 195 | | 200 | | 205 | |
| Ser Val Ser Arg Asn Ser Thr Val Arg Tyr Phe Asp Asn Gly Thr Ala | | | | | |
| 210 | | 215 | | 220 | |
| Leu Val Val Gln Trp Asp His Val His Leu Gln Asp Asn Tyr Asn Leu | | | | | |
| 225 | | 230 | | 235 | 240 |
| Gly Ser Phe Thr Phe Gln Ala Thr Leu Leu Met Asp Gly Arg Ile Ile | | | | | |
| | 245 | | 250 | | 255 |
| Phe Gly Tyr Lys Glu Ile Pro Val Leu Val Thr Gln Ile Ser Ser Thr | | | | | |
| | 260 | | 265 | | 270 |
| Asn His Pro Val Lys Val Gly Leu Ser Asp Ala Phe Val Val Val His | | | | | |
| | 275 | | 280 | | 285 |
| Arg Ile | | | | | |
| 290 | | | | | |

<210> 527

<211> 5343

<212> DNA

<213> Homo sapiens

<400> 527

```

nngtgccgtg tgctcctcac attcacgcag actgagactg agctgcccga ggaagagtgt
60
gaaggcccca agctgcccac cgaacggccc tgcttcttgg aagcatgcga tgagagcccg
120
gcctcccgag agctagacat ccctctccct gaggacagtg agacggctta cgactgggag
180
tacgctgggt tcacccttg cacagcaaca tgcttgggag gccatcaaga agccatagca
240
gtgtgcttac atatccagac ccagcagaca gtcaatgaca gcttgtgtga tatggtccac
300
cgctctccag ccatgagcca ggctgtgaac acagagccct gtccccccag gtggcatgtg
360
ggctcttggg ggccctgtc agctacctgt ggagttggaa ttcagaccg agatgtgtac
420
tgctgcacc caggggagac ccctgcccct cctgaggagt gccgagatga aaagccccat
480
gctttacaag catgcaatca gtttgactgc cctcctggct ggcacattga agaatggcag
540
cagtgttcca ggacttgtgg cgggggaact cagaacagaa gagtcacctg tcggcagctg
600
ctaacggatg gcagcttttt gaatctctca gatgaattgt gccaaaggacc caaggcatcg
660
tctcacaagt cctgtgccag gacagactgt cctccacatt tagctgtggg agactggtcg
720
aagtgttctg tcagttgtgg tggtggaatc cagagaagaa agcaggtgtg tcaaaggctg
780
gcagccaaag gtcggcgcac cccctcagc gagatgatgt gcagggatct accagggtc
840
cctcttgtaa gatcttgcca gatgcctgag tgcagtaaaa tcaaatcaga gatgaagaca
900
aaacttggtg agcaggggtc gcagatcctc agtgtccaga gagtctacat tcagacaagg
960

```

gaagagaagc gtattaacct gaccattggt agcagagcct atttgctgcc caacacatcc
1020
gtgattatta agtgccccgt gcgacgattc cagaaatctc tgatccagtg ggagaaggat
1080
ggccgttgcc tgcagaactc caaacggctt ggcatcacca agtcaggctc actaaaaatc
1140
cacggctctg ctgccccga catcggcggtg taccgggtgca ttgcaggctc tgcacaggaa
1200
acagttgtgc tcaagctcat tggtagtgac aaccgggtca tcgcacgccc agccctcagg
1260
gagcctatga gggaatatcc tgggatggac cacagcgaag ccaatagttt gggagtcaca
1320
tggcacaaaa tgaggcaaat gtggaataac aaaaatgacc tttatctgga tgatgaccac
1380
attagtaacc agcctttctt gagagctctg ttaggccact gcagcaattc tgcaggaagc
1440
accaactcct gggagttgaa gaataagcag tttgaagcag cagttaaaca aggagcatat
1500
agcatggata cagcccagtt tgatgagctg ataagaaaca tgagtcagct catggaaacc
1560
ggagaggta gcatgatct tgcgtcccag ctgatatac agctggtggc cgaattagcc
1620
aaggcacagc caacacacat gcagtggcgg ggcatccagg aagagacacc tctgtctgt
1680
cagctcagag gggaaacagg gagtgtgtcc caaagctcg atgcaaaaaa ctcaggcaag
1740
ctgacattca agccgaaagg acctgttctc atgaggcaaa gccaacctcc ctcaatttca
1800
tttaataaaa caataaatc caggattgga aatacagtat acattacaaa aaggacagag
1860
gtcatcaata tactgtgtga cttattacc ccagtgagg ccacatatac atggaccaag
1920
gatggaacct tggtacagcc ctcatgaaa ataattttgg atggaactgg gaagatacag
1980
atacagaatc ctacaaggaa agaacaaggc atatatgaat gttctgtagc taatcatctt
2040
ggttcagatg tggaaagtcc ttctgtgtg tatgcagagg cacctgtcat cttgtctgtt
2100
gaaagaaata tcaccaaac agagcacaac catctgtctg ttgtggttg aggcatctg
2160
gaggcagccc ttggagcaaa cgtgacaatc cgatgtcctg taaaagggtg ccctcagcct
2220
aatataactt ggttgaagag aggaggatct ctgagtggca atgtttcctt gcttttcaat
2280
ggatccctgt tggtgcagaa tgtttccctt gaaaatgaag gaacctacgt ctgcatagcc
2340
accaatgctc ttggaaaggc agtggaaca tctgtactcc acttgctgga acgaagatgg
2400
ccagagagta gaatcgtatt tctgcaagga cataaaaagt acattctcca ggcaaccaac
2460
actagaacca acagcaatga cccaacagga gaacccccgc ctcaagagcc tttttggag
2520
cctggttaact ggtcacattg ttctgccacc tgtggtcatt tgggagcccg cattcagaga
2580

ccccagtgtg tgatggccaa tgggcaggaa gtgagtgagg ccctgtgtga tcagcctcca
2640
gaagccactg gctgggtttg agccctgtaa catccgggac tgcccagcga ggtgggtcac
2700
aagtgtgtgg tcacagtgtc ctgtgtcttg cgggtgaagga taccacagtc ggcaggtgac
2760
gtgcaagcgg acaaaagcca atggaactgt gcaggtggtg tctccaagag catgtgcccc
2820
taaagaccgg cctctgggaa gaaaaccatg ttttgggtcat ccatgtgttc agtgggaacc
2880
aggggaaccg tgctctggac gttgcatggg ccgtgctgtg aggatgcagc agcgtcacac
2940
agcttgtcaa cacaacagct ctgactccaa ctgtgatgac agaaagagac ccaccttaag
3000
aagggaactg acatcagggg cctgtgatgt gtgttggcac acaggccctt ggaagccctg
3060
tacagcagcc tgtggcaggg gtttccagtc tcggaaaagtc gactgtatcc acacaaggag
3120
ttgcaaacct gtggccaaga gacactgtgt acagaaaaag aaaccaatct cctggcggca
3180
ctgtcttggg ccctcctgtg atagagactg cacagacaca actcactact gtatgtttgt
3240
aaaacatctt aatttgtgtt ctctagaccg ctacaaacaa aggtgctgcc agtcatgtca
3300
agagggataa acctttggag gggtcatgat gctgctgtga agataaaagt agaataaaa
3360
agctcttttc cccatgtcgc tgattcaaaa acatgtatct cttaaaagac tagattctat
3420
ggatcaaaca gaggttgatg caaaaacacc actgttaagg tgtaaagtga aattttccaa
3480
tggtagtttt atattccaat tttttaaaat gatgtattca aggatgaaca aaatactata
3540
gcatgcatgc cactgcactt gggacctcat catgtcagtt gaatcgagaa atcaccaaga
3600
ttatgagtgc atcctcacgt gctgcctctt tcctgtgata tgtagactag cacagagtgg
3660
tacatcctaa aaacttggga aacacagcaa cccatgactt cctcttctct caagttgcag
3720
gttttcaaca gttttataag gtatttgcac tttagaagct ctggccagta gttgttaaga
3780
tgttggcatt aatggcattt tcatagatcc ttggtttagt ctgtgaaaaa gaaaccatct
3840
ctctggatag gctgtcacac tgactgacct aagggttcat ggaagcatgg catcttgtcc
3900
ttgcttttag aacacccatg gaagaaaaca cagagtagat attgctgtca tttatacaac
3960
tacagaaatt tatctatgac ctaatgaggc atctcggaag tcaaagaaga gggaaagtta
4020
accttttcta ctgatttcgt agtatattca gagctttctt ttaagagctg tgaatgaaac
4080
tttttctaag cactattcta ttgcacacaa acagaaaacc aaagccttat tagacctaat
4140
ttatgcataa agtagtatcc ctgagaactt tattttggaa aatttataag aaagtaatcc
4200

aaataagaaa cacgatatgt gaaaataatt tttatagtaa ataattgttt tgggctgatt
 4260
 tttcagtaaa tccaaagtga cttagggttag aagttacact aaggaccagg ggttggaaac
 4320
 agaatttagt ttaagatttg aggaaaaggg taagggttag tttcagtttt aggattagag
 4380
 ctagaattgg gttagggtgag aaagaaagt aaggttaagg ctagagttgt ctttaagggt
 4440
 taggggttag accagggttag gtcagggttg gattgggttt agattggggc cagtgtggt
 4500
 gttagtata gtgtcaggat ggagggttag tttggagtaa gcgtgtgtgc tgaagtgagt
 4560
 tcaggctagc attaaattgt aagttctgaa gctgatttgg ttatggggtc tttccctgt
 4620
 atactaccag ttgtgtcttt agatggcaca caagtccaaa taagtgggtca tacttcttta
 4680
 ttcagggtct cagctgcctg tacacctgct gcctacatct tcttggcaac aaagttacct
 4740
 gccacaggct ctgctgagcc tagttcctgg tcagtaataa ctgaacagtg cattttggct
 4800
 ttggatgtgt ctgtggacaa gcttgtgtgag tttctctacc atattctgag cacacggtct
 4860
 cttttgttct aacttcagct tcactgacac tgggttgagc actactgtat gtggaggggt
 4920
 tgggtgattgg gaatggatgg gggacagtga ggaggacaca ccagcccatt agttgttaat
 4980
 catcaatcac atctgattgt tgaaggttat taaattaaaa gaaagatcat ttgtaacata
 5040
 ctctttgtat atatttatta tatgaaagggt gcaatatttt atttgtaca gtatgtaata
 5100
 aagacatggg acatatattt ttcttattaa caaaatttca tattaaattg cttcactttg
 5160
 tattttaaagt taaaagttac tatttttcat ttgctattgt actttcattg ttgtcattca
 5220
 attgacattc ctgtgtactg tattttacta ctgtttttat aacatgagag ttaatgtttc
 5280
 tgtttcatga tccttatgta attcagaaat aaatttactt tgattattca gtggcatcct
 5340
 tat
 5343

<210> 528

<211> 886

<212> PRT

<213> Homo sapiens

<400> 528

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Xaa | Cys | Arg | Val | Leu | Leu | Thr | Phe | Thr | Gln | Thr | Glu | Thr | Glu | Leu | Pro |
| 1 | | | | 5 | | | | 10 | | | | | 15 | | |
| Glu | Glu | Glu | Cys | Glu | Gly | Pro | Lys | Leu | Pro | Thr | Glu | Arg | Pro | Cys | Phe |
| | | | 20 | | | | | 25 | | | | | 30 | | |
| Leu | Glu | Ala | Cys | Asp | Glu | Ser | Pro | Ala | Ser | Arg | Glu | Leu | Asp | Ile | Pro |
| | | | 35 | | | | 40 | | | | | 45 | | | |
| Leu | Pro | Glu | Asp | Ser | Glu | Thr | Ala | Tyr | Asp | Trp | Glu | Tyr | Ala | Gly | Phe |

| | | |
|---------------------|---------------------|-------------------------|
| 50 | 55 | 60 |
| Thr Pro Cys Thr Ala | Thr Cys Leu Gly Gly | His Gln Glu Ala Ile Ala |
| 65 | 70 | 75 |
| Val Cys Leu His Ile | Gln Thr Gln Gln Thr | Val Asn Asp Ser Leu Cys |
| 85 | 90 | 95 |
| Asp Met Val His Arg | Pro Pro Ala Met Ser | Gln Ala Cys Asn Thr Glu |
| 100 | 105 | 110 |
| Pro Cys Pro Pro Arg | Trp His Val Gly Ser | Trp Gly Pro Cys Ser Ala |
| 115 | 120 | 125 |
| Thr Cys Gly Val Gly | Ile Gln Thr Arg Asp | Val Tyr Cys Leu His Pro |
| 130 | 135 | 140 |
| Gly Glu Thr Pro Ala | Pro Pro Glu Glu Cys | Arg Asp Glu Lys Pro His |
| 145 | 150 | 155 |
| Ala Leu Gln Ala Cys | Asn Gln Phe Asp Cys | Pro Pro Gly Trp His Ile |
| 165 | 170 | 175 |
| Glu Glu Trp Gln Gln | Cys Ser Arg Thr Cys | Gly Gly Gly Thr Gln Asn |
| 180 | 185 | 190 |
| Arg Arg Val Thr Cys | Arg Gln Leu Leu Thr | Asp Gly Ser Phe Leu Asn |
| 195 | 200 | 205 |
| Leu Ser Asp Glu Leu | Cys Gln Gly Pro Lys | Ala Ser Ser His Lys Ser |
| 210 | 215 | 220 |
| Cys Ala Arg Thr Asp | Cys Pro Pro His Leu | Ala Val Gly Asp Trp Ser |
| 225 | 230 | 235 |
| Lys Cys Ser Val Ser | Cys Gly Val Gly Ile | Gln Arg Arg Lys Gln Val |
| 245 | 250 | 255 |
| Cys Gln Arg Leu Ala | Ala Lys Gly Arg Arg | Ile Pro Leu Ser Glu Met |
| 260 | 265 | 270 |
| Met Cys Arg Asp Leu | Pro Gly Leu Pro Leu | Val Arg Ser Cys Gln Met |
| 275 | 280 | 285 |
| Pro Glu Cys Ser Lys | Ile Lys Ser Glu Met | Lys Thr Lys Leu Gly Glu |
| 290 | 295 | 300 |
| Gln Gly Pro Gln Ile | Leu Ser Val Gln Arg | Val Tyr Ile Gln Thr Arg |
| 305 | 310 | 315 |
| Glu Glu Lys Arg Ile | Asn Leu Thr Ile Gly | Ser Arg Ala Tyr Leu Leu |
| 325 | 330 | 335 |
| Pro Asn Thr Ser Val | Ile Ile Lys Cys Pro | Val Arg Arg Phe Gln Lys |
| 340 | 345 | 350 |
| Ser Leu Ile Gln Trp | Glu Lys Asp Gly Arg | Cys Leu Gln Asn Ser Lys |
| 355 | 360 | 365 |
| Arg Leu Gly Ile Thr | Lys Ser Gly Ser Leu | Lys Ile His Gly Leu Ala |
| 370 | 375 | 380 |
| Ala Pro Asp Ile Gly | Val Tyr Arg Cys Ile | Ala Gly Ser Ala Gln Glu |
| 385 | 390 | 395 |
| Thr Val Val Leu Lys | Leu Ile Gly Thr Asp | Asn Arg Leu Ile Ala Arg |
| 405 | 410 | 415 |
| Pro Ala Leu Arg Glu | Pro Met Arg Glu Tyr | Pro Gly Met Asp His Ser |
| 420 | 425 | 430 |
| Glu Ala Asn Ser Leu | Gly Val Thr Trp His | Lys Met Arg Gln Met Trp |
| 435 | 440 | 445 |
| Asn Asn Lys Asn Asp | Leu Tyr Leu Asp Asp | Asp His Ile Ser Asn Gln |
| 450 | 455 | 460 |
| Pro Phe Leu Arg Ala | Leu Leu Gly His Cys | Ser Asn Ser Ala Gly Ser |
| 465 | 470 | 475 |
| Thr Asn Ser Trp Glu | Leu Lys Asn Lys Gln | Phe Glu Ala Ala Val Lys |

<210> 529
<211> 4566

<212> DNA

<213> Homo sapiens

<400> 529

nggcgagcta agccggagga tgtgcagctg cggcggcggc gccggctacg aagaggacgg
60
ggacaggcgc cgtgcgaacc gagcccagcc agccggagga cgcgggcagg gcgggacggg
120
agcccggact cgtctgccgc cgcctcgtc gccgtcgtgc cggccccgcg tccccgcgcg
180
cgagcgggag gagccgccgc cacctcgcgc ccgagccgcc gctagcgcgc gccgggcatg
240
gtccccctctt aaaggcgag gccgcggcgg cgggggcggg cgtgcggaac aaagcgccgg
300
cgcggggcct gcgggcggct cgggggcgcg gatgggcgcg gcgggcccgc ggcggcggcg
360
gcgctgcccg ggccgggcct cgcggcgcta gggcgggctg gcctccgcgg gcgggggcag
420
cgggctgagg gcgcgcgggg cctgcggcgg cggcggcggc ggcggcggcg gcccgcgggg
480
cggagcggcg cgggcatggc cgcgcgcggc cggcgcgcct ggctcagcgt gctgctcggg
540
ctcgtcctgg gcttcgtgct ggctcgcgg ctcgtcctgc cccgggcttc cgagctgaag
600
cgagcgggcc cacggcgccg cgcagcccc gagggctgcc ggtccgggca ggcggcggct
660
tccagggcgc gcggggcgcg cggcgatgcg cgcggggcgc agctctggcc gcccggtcgc
720
gaccagatg gcggccccgc cgacaggaac tttctcttcg tgggagtcac gaccgccag
780
aaatacctgc agactcgggc cgtggccgcc tacagaacat ggtccaagac aattcctggg
840
aaagttcagt tcttctcaag tgagggttct gacacatctg taccaattcc agtagtgcca
900
ctacggggtg tggacgactc ctaccgccc cagaagaagt cttcatgat gctcaagtac
960
atgcacgacc actacttggc caagtatgaa tggtttatga gagcagatga tgacgtgtac
1020
atcaaaggag accgtctgga gaacttctg aggagtttga acagcagcga gcccctcttt
1080
cttgggcaga caggcctggg caccacggaa gaaatgggaa aactggccct ggagcctggt
1140
gagaacttct gcatgggggg gcctggcgtg atcatgagcc gggagggtgct tcggagaatg
1200
gtgccgcaca ttggcaagtg tctccgggag atgtacacca cccatgagga cgtggagggtg
1260
ggaaggtgtg tccggagggt tgcaggggtg cagtgtgtct ggtcttatga gatgcagcag
1320
cttttttatg agaattacga gcagaacaaa aaggggtaca ttagagatct ccataacagt
1380
aaaattcacc aagctatcac attacacccc aacaaaaacc caccctacca gtacaggctc
1440
cacagctaca tgctgagccg caagatatcc gagctccgcc atcgacaaat acagctgcac
1500

cgcgaaattg tectgatgag caaatacagc aacacagaaa ttcataaaga ggacctccag
1560
ctgggaatcc ctccctcctt catgaggttt cagccccgcc agcgagagga gattctggaa
1620
tgggagtttc tgactggaaa atacttgat tccggcagttg acggccagcc ccctcgaaga
1680
ggaatggact ccgcccagag ggaagccttg gacgacattg tcatgcaggt catggagatg
1740
atcaatgcca acgccaagac cagagggcgc atcattgact tcaaagagat ccagtacggc
1800
taccgccggg tgaaccccat gtatggggct gagtacatcc tggacctgct gcttctgtac
1860
aaaaagcaca aagggaagaa aatgacggtc cctgtgagga ggcacgcgta ttacagcag
1920
actttcagca aaatccagtt tgtggagcat gaggagctgg atgcacaaga gttggccaag
1980
agaatcaatc aggaatctgg atccttgctc tttctctcaa actccctgaa gaagctcgtc
2040
ccctttcagc tccctgggct gaagagtggc cacaagaac ccaaagataa aaagataaac
2100
atactgattc ctttgtctgg gcgtttcgac atgtttgtga gatttatggg aaactttgag
2160
aagacgtgct ttatcccaa tcagaacgct aagctcgagg ttctgctttt caattctgac
2220
tccaacctg acaaggccaa acaagttgaa ctgatgacag attaccgcat taagtacct
2280
aaagccgaca tgcagatttt gcctgtgtct ggagagtttt caagagccct ggccctggaa
2340
gtaggatcct ccagtttaa caatgaatct ttgctcttct tctgcgacgt cgacctcgtg
2400
tttactacag aattccttca gcgatgtcga gcaaatacag ttctgggcca acaaatatat
2460
tttccaatca tcttcagcca gtatgaccca aagattgttt atagtgggaa agttcccagt
2520
gacaaccatt ttgcctttac tcagaaaact ggcttctgga gaaactatgg gtttggcatc
2580
acgtgtattt ataagggaga tcttgtccga gtgggtggct ttgatgtttc catccaaggc
2640
tgggggctgg aggatgtgga ctttttcaac aaggttgtcc aggcaggttt gaagacgttt
2700
aggagccagg aagtaggagt agtcacgct caccatcctg tcttttgtga tccaatctt
2760
gaccccaaac agtacaaaat gtgcttgggg tccaaagcat cgacctatgg gtccacacag
2820
cagctggctg agatgtggct ggaaaaaat gatccaagtt acagtaaaag cagcaataat
2880
aatggctcag tgaggacagc ctaatgtcca gctttgctgg aaaagacgtt ttaattatc
2940
taatttattt ttcaaaaatt ttttgtatga tcagtttttg aagtccgtat acaaggatat
3000
attttacaag tggttttctt acataggact cttttaagat tgagctttct gaacaagaag
3060
gtgatcagtg tttgcctttg aacacatctt cttgctgaac attatgtagc agacctgctt
3120

aactttgact tgaaatgtac ctgatgaaca aaactttttt aaaaaaatgt tttcttttga
3180
gaccctttgc tccagtccta tggcagaaaa cgtgaacatt cctgcaaagt attattgtaa
3240
caaaacactg taactctggg aaatgttctg ttgtgattgt taacattcca cagattctac
3300
cttttgggtt ttgttttttt ttttttacia ttgtttttaa gccatttcat gttccagttg
3360
taagataagg aaatgtgata atagctgttt catcattgtc ttcaggagag ctttccagag
3420
ttgatcattt cccctcatgg tactctgctc agcatggcca cgtagggttt ttgtttgttt
3480
tgttttgttc tttttttgag acggagtctc actctgttac ccaggctgga atgcagtggc
3540
gcaatcttgg ctcaatttaa cctccacttc cctgggtcaa gcaattcccc tgcctttgcc
3600
tcccagtag ctgggattac aggcacacac caccacgcc agctagtttt tttgtatttt
3660
tagtagagac ggggtttcac catgcaagcc cagctggcca cgtagggttt aaagcaaggg
3720
gcgtgaagaa ggcacagtga ggtatgtggc tgttctcgtg gtagttcatt cggcctaaat
3780
agacctggca ttaaatttca agaaggattt ggcattttct cttcttgacc cttctcttta
3840
aagggtaaaa tattaatgtt tagaatgaca aagatgaatt attacaataa atctgatgta
3900
cacagactga aacacacaca catacaccct aatcaaaacg ttggggaaaa atgtatttgg
3960
ttttgttctt ttcactctgt ctgtgttatg tgggtggaga tgggtttcat tctttcatta
4020
ctgttttgtt ttatcctttg tatctgaaat acctttaatt tatttaatat ctgttggtca
4080
gagctctgcc atttcttgag tacctgttag ttagtattat ttatgtgtat cgggagtgtg
4140
tttagtctgt tttatttgca gtaaaccgat ctccaaagat ttccttttgg aaacgctttt
4200
tcccctcctt aatttttata ttccttactg ttttactaaa tattaagtgt tctttgacaa
4260
ttttggtgct catgtgtttt ggggacaaaa gtgaaatgaa tctgtcatta taccagaaag
4320
ttaaattctc agatcaaagt tgccttaata aatttgtttt catttagatt tcaaacagtg
4380
atagacttgc cattttaata cacgtcattg gagggctgag tatttgtaa tagcctgatg
4440
ctcatttggg aaaataaacc agtgaacaat atttttctat tgtacttttc gaaccatttt
4500
gtctcattat tcctgtttta gctgaagaat tgtattacat ttggagagta aaaaacttaa
4560
acacga
4566

<210> 530

<211> 802

<212> PRT

<213> Homo sapiens

<400> 530

```

Met Ala Ala Arg Gly Arg Arg Ala Trp Leu Ser Val Leu Leu Gly Leu
 1           5           10           15
Val Leu Gly Phe Val Leu Ala Ser Arg Leu Val Leu Pro Arg Ala Ser
 20           25           30
Glu Leu Lys Arg Ala Gly Pro Arg Arg Arg Ala Ser Pro Glu Gly Cys
 35           40           45
Arg Ser Gly Gln Ala Ala Ala Ser Gln Ala Gly Gly Ala Arg Gly Asp
 50           55           60
Ala Arg Gly Ala Gln Leu Trp Pro Pro Gly Ser Asp Pro Asp Gly Gly
 65           70           75           80
Pro Arg Asp Arg Asn Phe Leu Phe Val Gly Val Met Thr Ala Gln Lys
 85           90           95
Tyr Leu Gln Thr Arg Ala Val Ala Ala Tyr Arg Thr Trp Ser Lys Thr
100           105           110
Ile Pro Gly Lys Val Gln Phe Phe Ser Ser Glu Gly Ser Asp Thr Ser
115           120           125
Val Pro Ile Pro Val Val Pro Leu Arg Gly Val Asp Asp Ser Tyr Pro
130           135           140
Pro Gln Lys Lys Ser Phe Met Met Leu Lys Tyr Met His Asp His Tyr
145           150           155           160
Leu Asp Lys Tyr Glu Trp Phe Met Arg Ala Asp Asp Asp Val Tyr Ile
165           170           175
Lys Gly Asp Arg Leu Glu Asn Phe Leu Arg Ser Leu Asn Ser Ser Glu
180           185           190
Pro Leu Phe Leu Gly Gln Thr Gly Leu Gly Thr Thr Glu Glu Met Gly
195           200           205
Lys Leu Ala Leu Glu Pro Gly Glu Asn Phe Cys Met Gly Gly Pro Gly
210           215           220
Val Ile Met Ser Arg Glu Val Leu Arg Arg Met Val Pro His Ile Gly
225           230           235           240
Lys Cys Leu Arg Glu Met Tyr Thr Thr His Glu Asp Val Glu Val Gly
245           250           255
Arg Cys Val Arg Arg Phe Ala Gly Val Gln Cys Val Trp Ser Tyr Glu
260           265           270
Met Gln Gln Leu Phe Tyr Glu Asn Tyr Glu Gln Asn Lys Lys Gly Tyr
275           280           285
Ile Arg Asp Leu His Asn Ser Lys Ile His Gln Ala Ile Thr Leu His
290           295           300
Pro Asn Lys Asn Pro Pro Tyr Gln Tyr Arg Leu His Ser Tyr Met Leu
305           310           315           320
Ser Arg Lys Ile Ser Glu Leu Arg His Arg Thr Ile Gln Leu His Arg
325           330           335
Glu Ile Val Leu Met Ser Lys Tyr Ser Asn Thr Glu Ile His Lys Glu
340           345           350
Asp Leu Gln Leu Gly Ile Pro Pro Ser Phe Met Arg Phe Gln Pro Arg
355           360           365
Gln Arg Glu Glu Ile Leu Glu Trp Glu Phe Leu Thr Gly Lys Tyr Leu
370           375           380
Tyr Ser Ala Val Asp Gly Gln Pro Pro Arg Arg Gly Met Asp Ser Ala
385           390           395           400
Gln Arg Glu Ala Leu Asp Asp Ile Val Met Gln Val Met Glu Met Ile

```

```

      405      410      415
Asn Ala Asn Ala Lys Thr Arg Gly Arg Ile Ile Asp Phe Lys Glu Ile
      420      425      430
Gln Tyr Gly Tyr Arg Arg Val Asn Pro Met Tyr Gly Ala Glu Tyr Ile
      435      440      445
Leu Asp Leu Leu Leu Leu Tyr Lys Lys His Lys Gly Lys Lys Met Thr
      450      455      460
Val Pro Val Arg Arg His Ala Tyr Leu Gln Gln Thr Phe Ser Lys Ile
465      470      475      480
Gln Phe Val Glu His Glu Glu Leu Asp Ala Gln Glu Leu Ala Lys Arg
      485      490      495
Ile Asn Gln Glu Ser Gly Ser Leu Ser Phe Leu Ser Asn Ser Leu Lys
      500      505      510
Lys Leu Val Pro Phe Gln Leu Pro Gly Ser Lys Ser Glu His Lys Glu
      515      520      525
Pro Lys Asp Lys Lys Ile Asn Ile Leu Ile Pro Leu Ser Gly Arg Phe
      530      535      540
Asp Met Phe Val Arg Phe Met Gly Asn Phe Glu Lys Thr Cys Leu Ile
545      550      555      560
Pro Asn Gln Asn Val Lys Leu Val Val Leu Leu Phe Asn Ser Asp Ser
      565      570      575
Asn Pro Asp Lys Ala Lys Gln Val Glu Leu Met Thr Asp Tyr Arg Ile
      580      585      590
Lys Tyr Pro Lys Ala Asp Met Gln Ile Leu Pro Val Ser Gly Glu Phe
      595      600      605
Ser Arg Ala Leu Ala Leu Glu Val Gly Ser Ser Gln Phe Asn Asn Glu
      610      615      620
Ser Leu Leu Phe Phe Cys Asp Val Asp Leu Val Phe Thr Thr Glu Phe
625      630      635      640
Leu Gln Arg Cys Arg Ala Asn Thr Val Leu Gly Gln Gln Ile Tyr Phe
      645      650      655
Pro Ile Ile Phe Ser Gln Tyr Asp Pro Lys Ile Val Tyr Ser Gly Lys
      660      665      670
Val Pro Ser Asp Asn His Phe Ala Phe Thr Gln Lys Thr Gly Phe Trp
      675      680      685
Arg Asn Tyr Gly Phe Gly Ile Thr Cys Ile Tyr Lys Gly Asp Leu Val
      690      695      700
Arg Val Gly Gly Phe Asp Val Ser Ile Gln Gly Trp Gly Leu Glu Asp
705      710      715      720
Val Asp Leu Phe Asn Lys Val Val Gln Ala Gly Leu Lys Thr Phe Arg
      725      730      735
Ser Gln Glu Val Gly Val Val His Val His His Pro Val Phe Cys Asp
      740      745      750
Pro Asn Leu Asp Pro Lys Gln Tyr Lys Met Cys Leu Gly Ser Lys Ala
      755      760      765
Ser Thr Tyr Gly Ser Thr Gln Gln Leu Ala Glu Met Trp Leu Glu Lys
      770      775      780
Asn Asp Pro Ser Tyr Ser Lys Ser Ser Asn Asn Asn Gly Ser Val Arg
785      790      795      800
Thr Ala

```

<210> 531

<211> 321

<212> DNA

<213> Homo sapiens

<400> 531

ngatgatgaa tccccccgca gcctcgtaa tatggggggc ttcctacccc agcaaaaggc
 60
 acggcaatac gtctcgaaca aaggtctttt gtttcgaaat aacaaggggt tagagctaag
 120
 aggaagaagc gtgaaacgct gtaggaccag cgtttcgaac gccccgagg tgaaccctcg
 180
 ggggcgtctg aatcaggcca gttgggcctg ggacgacagc gggtgcagcg gcagcaatgg
 240
 cgcgtgcgga tcagccttga tcgattcacg ccaggcgccg agccactcgg cgtggccttc
 300
 gttccacacc tgctgggtgca g
 321

<210> 532

<211> 96

<212> PRT

<213> Homo sapiens

<400> 532

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Gly | Gly | Phe | Leu | Pro | Gln | Gln | Lys | Ala | Arg | Gln | Tyr | Val | Ser | Asn |
| 1 | | | | 5 | | | | 10 | | | | | 15 | | |
| Lys | Gly | Leu | Leu | Phe | Arg | Asn | Asn | Lys | Gly | Leu | Glu | Leu | Arg | Gly | Arg |
| | | 20 | | | | | 25 | | | | 30 | | | | |
| Ser | Val | Lys | Arg | Cys | Arg | Thr | Ser | Val | Ser | Asn | Ala | Pro | Glu | Val | Asn |
| | | 35 | | | | 40 | | | | | 45 | | | | |
| Pro | Arg | Gly | Arg | Leu | Asn | Gln | Ala | Ser | Trp | Ala | Trp | Asp | Asp | Ser | Gly |
| | 50 | | | | 55 | | | | 60 | | | | | | |
| Cys | Ser | Gly | Ser | Asn | Gly | Ala | Cys | Gly | Ser | Ala | Leu | Ile | Asp | Ser | Arg |
| 65 | | | | 70 | | | | 75 | | | | | 80 | | |
| Gln | Ala | Pro | Ser | His | Ser | Ala | Trp | Pro | Ser | Phe | His | Thr | Cys | Trp | Cys |
| | | | 85 | | | | | 90 | | | | | 95 | | |

<210> 533

<211> 335

<212> DNA

<213> Homo sapiens

<400> 533

nagtttccgg tgaaccgctc cgcaatgcct cgtgacatcg acttcagcga agccaacagg
 60
 agcatcatcg acaacatggc aactgcctca atcccgtttt tccgaaccga caaaaactgg
 120
 gagacgtggt cgagtcaggt ccggcatttc attagccttt tacacccaaa agtcaccctc
 180
 accaacattg acaacgtcct caacaaagat cacctgcgtt ggctacactt tcttttggag
 240
 ggctgcctgg agccaaacgt ggcctgatt gtccagggt actgttcgcc tggcaagctg
 300
 taccgcaagc ttgaggagct atatgccct tctgc
 335

<210> 534
 <211> 103
 <212> PRT
 <213> Homo sapiens

<400> 534
 Met Pro Arg Asp Ile Asp Phe Ser Glu Ala Asn Arg Ser Ile Ile Asp
 1 5 10 15
 Asn Met Ala Thr Ala Ser Ile Pro Leu Phe Arg Thr His Lys Asn Trp
 20 25 30
 Glu Thr Trp Ser Ser Gln Val Arg His Phe Ile Ser Leu Leu His Pro
 35 40 45
 Lys Val Thr Leu Thr Asn Ile Asp Asn Val Leu Asn Lys Asp His Leu
 50 55 60
 Arg Trp Leu His Phe Leu Leu Glu Gly Arg Leu Glu Pro Asn Val Arg
 65 70 75 80
 Leu Ile Val Gln Gly Tyr Cys Ser Pro Gly Lys Leu Tyr Arg Lys Leu
 85 90 95
 Glu Glu Leu Tyr Ala Pro Ser
 100

<210> 535
 <211> 402
 <212> DNA
 <213> Homo sapiens

<400> 535
 acgcgtctct acagccggac taagcacagg ctacgccccg gtcgccatgc gccaggctc
 60
 gggtatcagc cgaggaatcc acggcgaaat gaccagtagc ggccctaata caactatgct
 120
 gccgagcagc agacgtcgag gtcgggtcat gaggatgccg acggccaccg cgaccgggta
 180
 taccacaat gcaggaacaa ggctgatagc tagggctgac cacagagcca ggccgcctgc
 240
 cgaggaaacg cccccacct ggtgactgcc agtatcagca ccgcgcagct caacgacgtc
 300
 aacagtctcg ggattgacca accgccacgt atgcagggcc atgtggggga gaatcacccc
 360
 caacgccaat gctgtcaccg agcctcgggc taggccgccg gc
 402

<210> 536
 <211> 114
 <212> PRT
 <213> Homo sapiens

<400> 536
 Met Ala Leu His Thr Trp Arg Leu Val Asn Pro Glu Thr Val Asp Val
 1 5 10 15
 Val Glu Leu Arg Gly Ala Asp Thr Gly Ser His Gln Val Gly Gly Val
 20 25 30
 Ser Ser Ala Gly Gly Leu Ala Leu Trp Ser Ala Leu Ala Ile Ser Leu

```

      35      40      45
Val Pro Ala Leu Trp Val Tyr Pro Val Ala Val Ala Val Gly Ile Leu
  50      55      60
Met Thr Arg Pro Arg Arg Leu Leu Leu Gly Ser Ile Val Val Leu Gly
  65      70      75      80
Pro Leu Leu Val Ile Ser Pro Trp Ile Pro Arg Leu Ile Thr Glu Pro
      85      90      95
Gly Arg Met Ala Thr Gly Ala Glu Pro Val Leu Ser Pro Ala Val Glu
  100      105      110
Thr Arg

```

<210> 537
 <211> 404
 <212> DNA
 <213> Homo sapiens

```

<400> 537
gtgcacatcg gcggcaccga cttcgacaaa caactctcgc tggctggcat gatgccgctg
  60
ttcggctacg gcagccgcat gaagagcggc gcctacatgc ccaccagcca ccacatgaac
  120
ctggcgacct ggcacacccat caactcgggtg tactcgcaaa aatcccagct ggccctgggc
  180
agcatgcgct acgacatcga agacaccggc ggcacgcacc gcctgttcaa gctgatcgaa
  240
cagcgtgctg ggcactggct tgccatggaa gtggaagaaa ccaagatcca gtcacccat
  300
caagacagcc gccacgtgcc gctggaccgc atcgaagcgg gcctgagcgt agacctgagc
  360
cgggcgctgt tcgaatcgtc catcgacaac ctgctcgaac gcgt
  404

```

<210> 538
 <211> 118
 <212> PRT
 <213> Homo sapiens

```

<400> 538
Met Met Pro Leu Phe Gly Tyr Gly Ser Arg Met Lys Ser Gly Ala Tyr
  1      5      10      15
Met Pro Thr Ser His His Met Asn Leu Ala Thr Trp His Thr Ile Asn
      20      25      30
Ser Val Tyr Ser Gln Lys Ser Gln Leu Ala Leu Gly Ser Met Arg Tyr
      35      40      45
Asp Ile Glu Asp Thr Gly Gly Ile Asp Arg Leu Phe Lys Leu Ile Glu
      50      55      60
Gln Arg Ala Gly His Trp Leu Ala Met Glu Val Glu Glu Thr Lys Ile
      65      70      75      80
Gln Leu Thr His Gln Asp Ser Arg His Val Pro Leu Asp Arg Ile Glu
      85      90      95
Ala Gly Leu Ser Val Asp Leu Ser Arg Ala Leu Phe Glu Ser Ser Ile
      100      105      110
Asp Asn Leu Leu Glu Arg

```

115

<210> 539
 <211> 534
 <212> DNA
 <213> Homo sapiens

<400> 539
 nnacgcgtga aaaagaagaa aatgaaggaa agcgaggctg acagcgagggt gaagcatcaa
 60
 ccaattttca taaaagaaag attgaagctt tttgaaatac tgaagaaaga ccatcagctc
 120
 ttacttgcca tttatggaaa aaagggggat acaagcaaca tcatcacagt aagagtggct
 180
 gatgggcaaa cagtgcagg ggaagtctgg aaaacaacgc cttaccaagt ggctgctgaa
 240
 attagtcagg aactggctga aagcacggta atagccaaag tcaatgggtga actgtgggac
 300
 ctggaccgcc cattggaagg ggactcttct ctagagctgc ttacatttga taatgaggaa
 360
 gctcaagctg tgagtatttt aaaaccagac agccaaactt tgggtagtta tgttgtaaac
 420
 tacattatat aagaggccac atattgaatt cacgaatggt gagttttttg ggggtttcta
 480
 agatttaaaa tttgattatt gatgtttaat aaatatttgc ctcatgaatg ttaa
 534

<210> 540
 <211> 143
 <212> PRT
 <213> Homo sapiens

<400> 540
 Xaa Arg Val Lys Lys Lys Met Lys Glu Ser Glu Ala Asp Ser Glu
 1 5 10 15
 Val Lys His Gln Pro Ile Phe Ile Lys Glu Arg Leu Lys Leu Phe Glu
 20 25 30
 Ile Leu Lys Lys Asp His Gln Leu Leu Leu Ala Ile Tyr Gly Lys Lys
 35 40 45
 Gly Asp Thr Ser Asn Ile Ile Thr Val Arg Val Ala Asp Gly Gln Thr
 50 55 60
 Val Gln Gly Glu Val Trp Lys Thr Thr Pro Tyr Gln Val Ala Ala Glu
 65 70 75 80
 Ile Ser Gln Glu Leu Ala Glu Ser Thr Val Ile Ala Lys Val Asn Gly
 85 90 95
 Glu Leu Trp Asp Leu Asp Arg Pro Leu Glu Gly Asp Ser Ser Leu Glu
 100 105 110
 Leu Leu Thr Phe Asp Asn Glu Glu Ala Gln Ala Val Ser Ile Leu Lys
 115 120 125
 Pro Asp Ser Gln Thr Leu Gly Ser Tyr Val Val Asn Tyr Ile Ile
 130 135 140

<210> 541
 <211> 551

<212> DNA

<213> Homo sapiens

<400> 541

```

ggtagcgagc tgcgcgtgtg gtatgcggcc ttctatgcca agaagatgga caagcccatg
60
ctgaagcagg ccggctctgg cgtccacgct gcaggcaccc cagaaaacag cgcccccggtg
120
gagtcggagc ccagccagtg ggcgtgtaaa gtgtgttctg ccaccttcct ggagctgcag
180
ctcctcaatg gtaaggagga cgtgtgggga gccccagttg taaaactcct gtgtcgattt
240
ctctctgact tacgctgtca cctgtctgcg gctgtcgggg gtgtcccaga ctttgcctcg
300
tctgccccat tgccccacaa tgtagtcgcc agaaccaagg ctttctcagg gtttaaagct
360
tctgggcagt cccgcttccc acccccgacc cctgcaggcc tcaactcctca ctctcctgg
420
ttgggaagtt gcatttcagc tgggcgcctt gactctggag cactggcagg ggccaggggc
480
caggagccag ccgtggcatg tggtgtgcac tcttgccctt gttgtctcta cttgacagcc
540
ccctcacgcg t
551

```

<210> 542

<211> 168

<212> PRT

<213> Homo sapiens

<400> 542

```

Met Asp Lys Pro Met Leu Lys Gln Ala Gly Ser Gly Val His Ala Ala
1      5      10      15
Gly Thr Pro Glu Asn Ser Ala Pro Val Glu Ser Glu Pro Ser Gln Trp
20     25     30
Ala Cys Lys Val Cys Ser Ala Thr Phe Leu Glu Leu Gln Leu Leu Asn
35     40     45
Gly Lys Glu Asp Val Trp Gly Ala Pro Val Val Lys Leu Leu Cys Arg
50     55     60
Phe Leu Ser Asp Leu Arg Cys His Leu Ser Ala Ala Val Gly Gly Val
65     70     75     80
Pro Asp Phe Val Leu Ser Ala Pro Leu Pro His Asn Val Val Ala Arg
85     90     95
Thr Lys Ala Phe Ser Gly Phe Lys Ala Ser Gly Gln Ser Arg Phe Pro
100    105    110
Pro Pro Thr Pro Ala Gly Leu Thr Pro His Ser Ser Trp Leu Gly Ser
115    120    125
Cys Ile Ser Ala Gly Arg Leu Asp Ser Gly Ala Leu Ala Gly Ala Arg
130    135    140
Gly Gln Glu Pro Ala Val Ala Cys Val Val His Ser Cys Leu Cys Cys
145    150    155    160
Leu Tyr Leu Thr Ala Pro Ser Arg
165

```

<210> 543
 <211> 349
 <212> DNA
 <213> Homo sapiens

<400> 543
 nnaaagccgg acatgaatac ccgcattgct ggcaaaactg tcctgaccat cattctggcc
 60
 gggggcaaag gcagccgcct ggccccgatg accgatcagg tggccaaacc agccgtgccg
 120
 tttatgggga cgtaccgcct gattgacttt tcgctgtcca acattgtcca cagcggcttg
 180
 caggacgtct ggatcattga gcaaaacctg ccccatagct taaacgagca cctggctggg
 240
 gggcgctcct gggatctgga ccgcaccgcg ggtggcctga aggtcatgcc gcccttttcc
 300
 ggccctgccg atgaggacgg tggcttttcc gaaggcaacg cacacgcgt
 349

<210> 544
 <211> 116
 <212> PRT
 <213> Homo sapiens

<400> 544
 Xaa Lys Pro Asp Met Asn Thr Arg Ile Ala Gly Lys Thr Val Leu Thr
 1 5 10 15
 Ile Ile Leu Ala Gly Gly Lys Gly Ser Arg Leu Ala Pro Met Thr Asp
 20 25 30
 Gln Val Ala Lys Pro Ala Val Pro Phe Met Gly Thr Tyr Arg Leu Ile
 35 40 45
 Asp Phe Ser Leu Ser Asn Ile Val His Ser Gly Leu Gln Asp Val Trp
 50 55 60
 Ile Ile Glu Gln Asn Leu Pro His Ser Leu Asn Glu His Leu Ala Gly
 65 70 75 80
 Gly Arg Ser Trp Asp Leu Asp Arg Thr Arg Gly Gly Leu Lys Val Met
 85 90 95
 Pro Pro Phe Ser Gly Pro Ala Asp Glu Asp Gly Gly Phe Ser Glu Gly
 100 105 110
 Asn Ala His Ala
 115

<210> 545
 <211> 390
 <212> DNA
 <213> Homo sapiens

<400> 545
 catgatgcaa aaacagacat gcttatttca aaatataaaa gtgaaaaaga tcgttttagca
 60
 caagaaattg ttggtgtcat cacaggttct gcaatgccgg gtgggttcagc aaaccgtatc
 120
 ccaaataaag caggctcaaa tccagaaggt tctattgcaa cgcgttttat tgcagaaaca
 180

atgtataacg aactcaaaac agtggattta actattcaaa atgctggcgg tgtacgcgca
 240
 gatattttac cggggaatgt aacctttaac gatgcttata ctttcttacc ttctgggaat
 300
 acgttatata cctataaaat ggaaagtcca ttagtgaaac aagtgcctga agatgcaatg
 360
 ctatttgctt tgggtccccc ccccccccc
 390

<210> 546
 <211> 130
 <212> PRT
 <213> Homo sapiens

<400> 546
 His Asp Ala Lys Thr Asp Met Leu Ile Ser Lys Tyr Lys Ser Glu Lys
 1 5 10 15
 Asp Arg Leu Ala Gln Glu Ile Val Gly Val Ile Thr Gly Ser Ala Met
 20 25 30
 Pro Gly Gly Ser Ala Asn Arg Ile Pro Asn Lys Ala Gly Ser Asn Pro
 35 40 45
 Glu Gly Ser Ile Ala Thr Arg Phe Ile Ala Glu Thr Met Tyr Asn Glu
 50 55 60
 Leu Lys Thr Val Asp Leu Thr Ile Gln Asn Ala Gly Gly Val Arg Ala
 65 70 75 80
 Asp Ile Leu Pro Gly Asn Val Thr Phe Asn Asp Ala Tyr Thr Phe Leu
 85 90 95
 Pro Phe Gly Asn Thr Leu Tyr Thr Tyr Lys Met Glu Ser Ser Leu Val
 100 105 110
 Lys Gln Val Leu Glu Asp Ala Met Leu Phe Ala Leu Gly Pro Pro Pro
 115 120 125
 Pro Pro
 130

<210> 547
 <211> 306
 <212> DNA
 <213> Homo sapiens

<400> 547
 aagcttggtt ttctgatttt tattcaaatc tctatcatgg atgaagcatg cagtttcaga
 60
 atcagttcag tgttgacaac atatcaagat attctgcagt caatctcaat gtatgttcac
 120
 gaagcctcca acatattttg tgggatacca tctttgtcag gcattgtgct aggcactgtc
 180
 cctgcagtga ataagaaaga caggatttct gtatttatgg ggcttagtac caagttgttc
 240
 tcaaaacttc atgtttgtgt atacaaatca gctgaggcct tcaactaaact cnnnnnccnn
 300
 nncenn
 306

<210> 548